DARK MATTER IN COSMOLOGY
QUANTUM MEASUREMENTS
EXPERIMENTAL GRAVITATION

Matière Noire et Cosmologie - Mesures Quantiques

Gravitation Expérimentale

XVIth Moriond Worshop
Les Arcs, Savoie, France - January 20-27, 1996

Dark Matter in Cosmology - Quantum Measurements
Experimental Gravitation
Series : Moriond Workshops

ISBN 2-86332-203-6

EDITIONS FRONTIERES
B.P. 33
91192 Gif-sur-Yvette Cedex - France

Printed in Singapore

Proceedings of the XXXIst RENCONTRES DE MORIOND
Series : Moriond Workshops
Les Arcs, Savoie, France January 20-27, 1996

DARK MATTER IN COSMOLOGY
QUANTUM MEASUREMENTS
EXPERIMENTAL GRAVITATION

edited by

R. Ansari
Y. Giraud-Héraud
J. Trân Thanh Vân

EDITIONS
FRONTIERES

XVIth Moriond Workshop

Dark Matter in Cosmology - Quantum Measurements
Experimental Gravitation

was organized by

J. Trân Thanh Vân *(Orsay)*

with the active collaboration of :

E. Adelberger *(Seattle)*
R. Ansari *(Orsay)*
G. Chardin *(Saclay)*
T. Damour *(IHES, Bures/Yvette)*
O. Fackler *(Livermore)*
G. Fontaine *(Paris)*
G. Gerbier *(Saclay)*
Y. Giraud-Héraud *(Paris)*
G. Greene *(Los Alamos)*
B. Guiderdoni *(Paris)*
S. Haroche *(Paris)*
E. Hinds *(Cambridge)*
J. Kaplan *(Paris)*
J.-M. Raimond *(Paris)*
B. Sadoulet *(Berkeley)*

96' RENCONTRES DE MORIOND

The XXXIst Rencontres de Moriond were held in 1996 in Les Arcs 1800, Savoie, France.

The first meeting took place at Moriond in the French Alps in 1966. There, experimental as well as theoretical physicists not only shared their scientific preoccupations but also the household chores. The participants in the first meeting were mainly French physicists interested in electromagnetic interactions. In the following years, a session on high energy strong interactions was also added.

The main purpose of these meetings is to discuss recent developments in contemporary physics and also to promote effective collaboration between experimentalists and theorists in the field of elementary particle physics. By bringing together a relatively small number of participants, the meeting helps to develop better human relations as well as a more thorough and detailed discussion of the contributions.

This concern of research and experimentation of new channels of communication and dialogue which from the start animated the Moriond meetings, inspired us to organize a simultaneous meeting of biologists on Cell Differenciation (1970) and to create the Moriond Astrophysics Meeting (1981). In the same spirit, we have started a new series on Condensed Matter Physics in January 1994. Common meetings between biologists, astrophysicists, condensed matter physicists and high energy physicists are organized to study the implications of the advances from one field into the others. I hope that these conferences and lively discussions may give birth to new analytical methods or new mathematical languages.

At the XXXIst Rencontres de Moriond in 1996, four physics sessions, one astrophysics session and one biology session were held :

* January 20-27	"Dark matter in cosmology, Quantum measurements and Experimental gravitation.
	"Correlated fermions and transport in mesoscopic systems"
* March 16-23	"Electroweak Interactions and Unified Theories"
	"Microwave background Anisotropies"
* March 23-30	"QCD and High Energy Hadronic Interactions"
	" Rencontre de Biologie - Méribel "

I thank the organizers of the XXXIst Rencontres de Moriond :

- E. Adelberger, R. Ansari, G. Chardin, T. Damour, O. Fackler, G. Fontaine, G. Gerbier, Y. Giraud-Héraud, G. Greene , B. Guiderdoni, S. Haroche, E. Hinds, J. Kaplan, J-M. Raimond, B. Sadoulet for the session Dark matter in cosmology, Quantum measurements and Experimental gravitation,

- T. Jolicoeur, T. Martin, G. Montambaux, B. Pannetier for the session Correlated fermions and transport in mesoscopic systems

- P. Binetruy, A. Blondel, R. Cahn, G. Coignet, L. Fayard, P. Fayet, J.-M. Frère, H. Frisch, L. Krauss, L. Okun, S. Loucatos, C. Savoy and G. Unal for the Electroweak Interaction session,

- M. Bersanelli, F.R. Bouchet, M. Giard, R. Gispert, B. Guiderdoni, J.-M. Lamarre, N. Mandolesi, J.-L. Puget , J. Tauber for the Microwave background Anisotropies session,

- E. Augé, P. Aurenche, E. Berger, A. Capella, D. Denegri, B. Klima, L. Kluberg, L. McLerran, L. Montanet, B. Pietrzyck, D. Schiff, and C. Voltolini for the QCD and Hadronic Interaction session,

- F. Fabre, M. Fellous, P.-M. Guyon, I. Oliveri and K. Trân Thanh Vân for the Biology meeting.

and the conference secretaries : G. Ambonati, C. Andrews, N. Bachala, J.M. Ballocchi, C. Barthélémy, L. Besson, R. Bilhaut, C. Cohas-Bogey, Y. Gattignon, D. Larcher, D. Levaillant, H. Mauger, F. N. Osswald, , J. Raguideau, A. M. Renault-Perrin, J. Schmitt and F. Warin.

I am also grateful to Mr. E. Rocca-Serra, Ms. C. Dufourny, Ms. S. Kovalevsky who contributed through their hospitality and cooperation to the well-being of the participants enabling them to work in a relaxed atmosphere.

These Rencontres were sponsored by the European Union Human Capital and Mobility Programme, the Centre National de la Recherche Scientifique (INSU, SPM and FP), the Institut National de Physique Nucléaire et de Physique des Particules (IN2P3), the Commissariat à l'Energie Atomique (DAPNIA, DRECAM), the Ministère de l'Enseignement Supérieur et de la Recherche (programme ACCESS) and the International Science Foundation. I would like to express my thanks for their encouraging support.

I sincerely wish that a fruitful exchange and an efficient collaboration between the physicists, the astrophysicists and the biologists will arise from these Rencontres as from the previous ones.

J. Trân Thanh Vân

CONTENTS

III. QUANTUM MEASUREMENT

IV. EXPERIMENTAL GRAVITATION

DARK MATTER

WHY DO WE NEED NON-BARYONIC DARK MATTER?

Keith A. Olive
School of Physics and Astronomy, University of Minnesota, Minneapolis, MN 55455, USA

Abstract

Observational evidence along with theoretical arguments which call for non-baryonic dark matter are reviewed. A brief summary of the dark matter session is included.

There is increasing evidence that relative to the visible matter in the Universe, which is in the form of baryons, there is considerably more matter in the Universe that we don't see [1]. Here, I will review some of the motivations for dark matter in the Universe. The best observational evidence is found on the scale of galactic halos and comes from the observed flat rotation curves of galaxies. There is also mounting evidence for dark matter in elliptical galaxies as well as clusters of galaxies coming from X-ray observations of these objects. Also, direct evidence has been obtained through the study of gravitational lenses. In theory, we believe there is much more matter because 1) inflation tells us so (and there is at present no good alternative to inflation) and 2) our current understanding of galaxy formation only makes sense if there is more matter than we see. One can also make a strong case for the existence of non-baryonic dark matter in particular. The recurrent problem with baryonic dark matter is that not only is it very difficult to hide baryons, but the standard model of primordial nucleosynthesis would have to discarded if all of the dark matter is baryonic. Fortunately, as will be covered at length in these proceedings, there are several attractive alternatives to baryonic dark matter.

Before embarking on the subject of dark matter, it will be useful to review the relevant quantities from the standard big bang model. In a Friedmann-Robertson-Walker Universe, the expansion rate of the Universe (the Hubble parameter) is related to the energy density ρ and curvature constant k by

$$H^2 = \left(\frac{\dot{R}}{R}\right)^2 = \frac{8\pi G}{3}\rho - \frac{k}{R^2} \tag{1}$$

assuming no cosmological constant, where $k = \pm 1, 0$ for a closed, open or spatially flat Universe, and R is the cosmological scale factor. When $k = 0$, the energy density takes its "critical" value, $\rho = \rho_c = 3H^2/8\pi G = 1.88 \times 10^{-29} h_o^2$ g cm^{-3} where $h_o = H_o/100$ km s^{-1} Mpc^{-1} is the scaled present value of the Hubble parameter. The cosmological density parameter is defined by $\Omega = \rho/\rho_c$ and by rewriting eq. (1) we can relate k to Ω and H by

$$\frac{k}{R^2} = (\Omega - 1)H^2 \tag{2}$$

so that $k = +1, -1, 0$ corresponds to $\Omega > 1, < 1, = 1$.

In very broad terms, observational limits on the cosmological parameters are: $0.2 \lesssim \Omega \lesssim 2$ and $0.4 \lesssim h_o \lesssim 1.0$ [2]. The cosmological density is however sensitive to the particular scale being observed (at least on small scales). Accountably visible matter contributes in total only a small fraction to the overall density, giving $\Omega_V \sim .003 - .01$. In the bright central parts of galaxies, the density is larger $\Omega \sim 0.02 - 0.1$. On larger scales, that of binaries and small groups of galaxies, $\Omega \simeq 0.05 - 0.3$. On even larger scales the density may be large enough to support $\Omega \simeq 1.0$. Though there are no astronomical observations to support $\Omega > 1$, limits based on the deceleration of the Universe only indicate [2] that $\Omega \lesssim 2$.

The age of the Universe is also very sensitive to these parameters. Again, in the absence of a cosmological constant we have,

$$H_o t_U = \int_0^1 (1 - \Omega + \Omega/x)^{-1/2} dx \tag{3}$$

For $t_U > 13$ Gyr, $\Omega h_o^2 < 0.25$ if $h_o > 0.5$ and $\Omega h_o^2 < 0.45$ if $h_o > 0.4$. While for $t_U > 10$ Gyr, $\Omega h_o^2 < 0.8$ if $h_o > 0.5$ and $\Omega h_o^2 < 1.1$ if $h_o > 0.4$.

There is, in fact, good evidence for dark matter on the scale of galaxies (and their halos). Assuming that galaxies are in virial equilibrium, one expects that by Newton's Laws one can relate the mass at a given distance r, from the center of a galaxy to its rotational velocity

$$M(r) \propto v^2 r/G_N \tag{4}$$

The rotational velocity, v, is measured [3, 4, 5] by observing 21 cm emission lines in HI regions (neutral hydrogen) beyond the point where most of the light in the galaxy ceases. A compilation of nearly 1000 rotation curves of spiral galaxies have been plotted in [6] as a function of r for varying brightnesses. If the bulk of the mass is associated with light, then beyond the point where most of the light stops M would be constant and $v^2 \propto 1/r$. This is not the case, as the rotation curves appear to be flat, i.e., $v \sim$ constant outside the core of the galaxy. This implies that $M \propto r$ beyond the point where the light stops. This is one of the strongest pieces of evidence for the existence of dark matter. Velocity measurements indicate dark matter in elliptical galaxies as well [7].

Galactic rotation curves are not the only observational indication for the existence of dark matter. X-ray emitting hot gas in elliptical galaxies also provides an important piece of evidence for dark matter. As an example, consider the large elliptical M87. The detailed profiles of the temperature and density of the hot X-ray emitting gas have been mapped out [8]. By assuming hydrostatic equilibrium, these measurements allow one to determine the overall mass distribution in the galaxy necessary to bind the hot gas. Based on an isothermal model with temperature $kT = 3$keV (which leads to a conservative estimate of the total mass),

Fabricant and Gorenstein [8] predicted that the total mass out to a radial distance of 392 Kpc, is $5.7 \times 10^{13} M_{\odot}$ whereas the mass in the hot gas is only $2.8 \times 10^{12} M_{\odot}$ or only 5% of the total. The visible mass is expected to contribute only 1% of the total. The inferred value of Ω based on M87 would be ~ 0.2.

M87 is not the only example of an elliptical galaxy in which X-ray emitting hot gas is observed to indicate the presence of dark matter. At this meeting, Foreman [9], showed several examples of ellipiticals with large mass to light ratios. For example in the case of N4472, while the optical observations go out to 25 kpc, the X-ray gas is seen out to 75 kpc, indicating M/L's of about 60 at 70 kpc and up to 90 at 100 kpc. Similar inferences regarding the existence of dark matter can be made from the X-ray emission from small groups of galaxies [10, 11].

On very large scales, it is possible to get an estimate of Ω from the distribution of peculiar velocities. On scales, λ, where perturbations, δ, are still small, peculiar velocities can be expressed [12] as $v \sim H\lambda\delta\Omega^{0.6}$. On these scales, measurements of the peculiar velocity field from the IRAS galaxy catalogue indicate that indeed Ω is close to unity [13]. Another piece of evidence on large scales, is available from gravitational lensing [14]. The systematic lensing of the roughly 150,000 galaxies per deg^2 at redshifts between $z = 1 - 3$ into arcs and arclets allow one to trace the matter distribution in a foreground cluster. Van Waerbeke discussed recent results of weak gravitational lensing looking at systems, which if virialized, have mass to light ratios in the range 400–1000 and correspond to values of Ω between 0.25 and 0.6 [15]. Unfortunately, none of these observations reveal the identity of the dark matter.

Theoretically, there is no lack of support for the dark matter hypothesis. The standard big bang model including inflation almost requires that $\Omega = 1$ [16]. The simple and unfortunate fact that at present we do not even know whether Ω is larger or smaller than one indicates that we do not know the sign of the curvature term, further implying that it is subdominant in Eq. (1)

$$\frac{k}{R^2} < \frac{8\pi G}{3}\rho \tag{5}$$

In an adiabatically expanding Universe, $R \sim T^{-1}$ where T is the temperature of the thermal photon background. Therefore the quantity

$$\hat{k} = \frac{k}{R^2 T^2} < \frac{8\pi G}{3T_o^2} < 2 \times 10^{-58} \tag{6}$$

is dimensionless and constant in the standard model. The smallness of $\hat{k}$ is known as the curvature problem and can be resolved by a period of inflation. Before inflation, let us write $R = R_i$, $T = T_i$ and $R \sim T^{-1}$. During inflation, $R \sim T^{-1} \sim e^{Ht}$, where H is constant. After inflation, $R = R_f \gg R_i$ but $T = T_f = T_R \lesssim T_i$ where T_R is the temperature to which the Universe reheats. Thus $R \not\sim T^{-1}$ and $\hat{k} \to 0$ is not constant. But from Eqs. (2) and (6) if $\hat{k} \to 0$ then $\Omega \to 1$, and since typical inflationary models contain much more expansion than is necessary, Ω becomes exponentially close to one.

If this is the case and $\Omega = 1$, then we know two things: Dark matter exists, since we don't see $\Omega = 1$ in luminous objects, and most (at least 90%) of the dark matter is not baryonic. The latter conclusion is a result from big bang nucleosynthesis [17, 18], which constrains the baryon-to-photon ratio $\eta = n_B/n_\gamma$ to

$$1.4 \times 10^{-10} < \eta < 3.8 \times 10^{-10} \tag{7}$$

which corresponds to a limit on Ω_B

$$0.005 < \Omega_B < 0.09 \tag{8}$$

for $0.4 \lesssim h_o \lesssim 1.0$. Thus $1 - \Omega_B$ is not only dark but also non-baryonic. I will return to big bang nucleosynthesis below.

Another important piece of theoretical evidence for dark matter comes from the simple fact that we are here living in a galaxy. The type of perturbations produced by inflation [19] are, in most models, adiabatic perturbations ($\delta\rho/\rho \propto \delta T/T$), and I will restrict my attention to these. Indeed, the perturbations produced by inflation also have the very nearly scale-free spectrum described by Harrison and Zeldovich [20]. When produced, scale-free perturbations fall off as $\frac{\delta\rho}{\rho} \propto l^{-2}$ (increase as the square of the wave number). At early times $\delta\rho/\rho$ grows as t until the time when the horizon scale (which is proportional to the age of the Universe) is comparable to l. At later times, the growth halts (the mass contained within the volume l^3 has become smaller than the Jean's mass) and $\frac{\delta\rho}{\rho} = \delta$ (roughly) independent of the scale l. When the Universe becomes matter dominated, the Jean's mass drops dramatically and growth continues as $\frac{\delta\rho}{\rho} \propto R \sim 1/T$. For an overview of the evolution of density perturbations and the resulting observable spectrum see [21]. The transition to matter dominance is determined by setting the energy densities in radiation (photons and any massless neutrinos) equal to the energy density in matter (baryons and any dark matter). For three massless neutrinos and baryons (no dark matter), matter dominance begins at

$$T_m = 0.22 m_B \eta \tag{9}$$

and for $\eta < 3.8 \times 10^{-10}$, this corresponds to $T_m < 0.08$ eV.

The subsequent non-linear growth in $\delta\rho/\rho$ was discussed in these sessions in some detail at this meeting. In particular, there was a considerable discussion of the effects of the non-linear regime on the power spectrum and the appearance of non-Gaussian features such as skewness and kurtosis [22]. Colombi [23] discussed the 2- and 3-point correlation functions. Numerous simulations were presented to reflect the hydrodynamics of galaxy formation [24] and the origin of the large scale bias [25].

Because we are considering adiabatic perturbations, there will be anisotropies produced in the microwave background radiation on the order of $\delta T/T \sim \delta$. The value of δ, the amplitude of the density fluctuations at horizon crossing, has now been determined by COBE [26], $\delta = (5.7 \pm 1.4) \times 10^{-6}$. Without the existence of dark matter, $\delta\rho/\rho$ in baryons could then achieve a maximum value of only $\delta\rho/\rho \sim A_\lambda \delta(T_m/T_o) \lesssim 2 \times 10^{-3} A_\lambda$, where $T_o = 2.35 \times 10^{-4}$ eV is the present temperature of the microwave background and $A_\lambda \sim 1 - 10$ is a scale dependent growth factor. The overall growth in $\delta\rho/\rho$ is too small to argue that growth has entered a nonlinear regime needed to explain the large value (10^5) of $\delta\rho/\rho$ in galaxies.

Dark matter easily remedies this dilemma in the following way. The transition to matter dominance is determined by setting equal to each other the energy densities in radiation (photons and any massless neutrinos) and matter (baryons and any dark matter). While if we suppose that there exists dark matter with an abundance $Y_\chi = n_\chi/n_\gamma$ (the ratio of the number density of χ's to photons) then

$$T_m = 0.22 m_\chi Y_\chi \tag{10}$$

Since we can write $m_\chi Y_\chi/\mathrm{GeV} = \Omega_\chi h^2/(4 \times 10^7)$, we have $T_m/T_o = 2.4 \times 10^4 \Omega_\chi h^2$ which is to be compared with 350 in the case of baryons alone. The baryons, although still bound

to the radiation until decoupling, now see deep potential wells formed by the dark matter perturbations to fall into and are no longer required to grow at the rate $\delta\rho/\rho \propto R$.

With regard to dark matter and galaxy formation, all forms of dark matter are not equal. They can be distinguished by their relative temperature at T_m [27]. Particles which are still largely relativistic at T_m (like neutrinos or other particles with $m_\chi < 100$ eV) have the property [28] that (due to free streaming) they erase perturbations out to very large scales given by the Jean's mass

$$M_J = 3 \times 10^{18} \frac{M_\odot}{m_\nu{}^2(eV)} \tag{11}$$

Thus, very large scale structures form first and galaxies are expected to fragment out later. Particles with this property are termed hot dark matter particles. Cold particles ($m_\chi > 1$ MeV) have the opposite behavior. Small scale structure forms first aggregating to form larger structures later. Neither candidate is completely satisfactory when the resulting structure is compared to the observations. For more details, I refer the reader to reviews in refs. [1].

The most promising possibility we have for unscrambling the various possible scenarios for structure formation is the careful analysis of the observed power spectrum. Rewriting the density contrast in Fourier space,

$$\delta(k) \propto \int d^3x \frac{\delta\rho}{\rho}(x) e^{i\mathbf{k}\cdot x} \tag{12}$$

the power spectrum is just

$$P(k) = \langle |\delta(k)|^2 \rangle \tag{13}$$

which is often written in term of a transfer function and a power law, $P \sim T(k)k^n$. (The flat spectrum produced by inflation has $n = 1$.) The data contributing to $P(k)$ include observations of galaxy distributions, velocity distributions and of course the CMB [29]. However to make a comparison with our theoretical expectations, we rely on n-body simulations and a deconvolution of the theory from the observations. Overall, there is actually a considerable amount of concordance with our expectations. The velocity distributions indicate that $0.3 < \Omega < 1$ and the power spectrum is roughly consistent with an $\Omega = 1$, and $n = 1$, cold dark matter Universe.

Future surveys [30, 31] will however, most certainly dramatically improve our understanding of the detailed features of the power spectrum and their theoretical interpretations. We should in principle be able to distinguish between a mixed dark matter and a cold dark matter $\Omega = 1$ Universe, whether or not $\Omega < 1$, with CDM, the presence of a cosmological constant, or a tilt in the spectrum. Zurek [32], stressed the importance of numerical simulations in this context. Future probes of the small scale anisotropy [31, 33] should in addition be able to determine the values of not only Ω, but also Ω_B and h_o to a high degree of accuracy through a careful analysis of the Doppler peak in the power spectrum.

Accepting the dark matter hypothesis, the first choice for a candidate should be something we know to exist, baryons. Though baryonic dark matter can not be the whole story if $\Omega = 1$, the identity of the dark matter in galactic halos, which appear to contribute at the level of $\Omega \sim 0.05$, remains an important question needing to be resolved. A baryon density of this magnitude is not excluded by nucleosynthesis. Indeed we know some of the baryons are dark since $\Omega \lesssim 0.01$ in the disk of the galaxy.

It is quite difficult, however, to hide large amounts of baryonic matter [34]. Sites for halo baryons that have been discussed include snowballs, which tend to sublimate, cold hydrogen

gas, which must be supported against collapse, and hot gas, which can be excluded by the X-ray background. Stellar objects (collectively called MACHOs for macroscopic compact halo objects) must either be so small ($M < 0.08\ M_\odot$) so as not to have begun nuclear burning or so massive so as to have undergone total gravitational collapse without the ejection of heavy elements. Intermediate mass stars are generally quite problematic because either they are expected to still reside on the main-sequence today and hence would be visible, or they would have produced an excess of heavy elements.

On the other hand, MACHOs are a candidate which are testable by the gravitational microlensing of stars in a neighboring galaxy such as the LMC [35]. By observing millions of stars and examining their intensity as a function of time, it may be possible to determine the presence of dark objects in our halo. It is expected that during a lensing event, a star's intensity will rise in an achromatic fashion over a period $\delta t \sim 3\sqrt{M/.001 M_\odot}$ days. Indeed, microlensing candidates have been found [36]. For low mass objects, those with $M < 0.1 M_\odot$, it appears however that the halo fraction of Machos is small, $\approx 0.19^{+.16}_{-.10}$ [37]. There have been some recent reports of events with longer duration [38] leading to the speculation of a white dwarf population in the halo. Though it is too early to determine the implications of these observations, they are very encouraging in that perhaps this issue can and will be decided.

The degree to which baryons can contribute to dark matter depends ultimately on the overall baryon contribution to Ω which is constrained by nucleosynthesis. Because of its importance regarding the issue of dark matter and in particular non-baryonic dark matter, I want to review the status of big bang nucleosynthesis.

Conditions for the synthesis of the light elements were attained in the early Universe at temperatures $T \lesssim 1$ MeV, corresponding to an age of about 1 second. Given a single input parameter, the baryon-to-photon ratio, η, the theory is capable of predicting the primordial abundances of D/H, ^{3}He/H, ^{4}He/H and ^{7}Li/H. The comparison of these predictions with the observational determination of the abundances of the light elements not only tests the theory but also fixes the value of η.

The dominant product of big bang nucleosynthesis is ^{4}He resulting in an abundance of close to 25 % by mass. In the standard model, the ^{4}He mass fraction depends only weakly on η. When we go beyond the standard model, the ^{4}He abundance is very sensitive to changes in the expansion rate which can be related to the effective number of neutrino flavors. Lesser amounts of the other light elements are produced: D and ^{3}He at the level of about 10^{-5} by number, and ^{7}Li at the level of 10^{-10} by number.

There is now a good collection of abundance information on the ^{4}He mass fraction, Y, O/H, and N/H in over 50 extragalactic HII (ionized hydrogen) regions [39, 40, 41]. The observation of the heavy elements is important as the helium mass fraction observed in these HII regions has been augmented by some stellar processing, the degree to which is given by the oxygen and nitrogen abundances. In an extensive study based on the data in [39, 40], it was found [42] that the data is well represented by a linear correlation for Y vs. O/H and Y vs. N/H. It is then expected that the primordial abundance of ^{4}He can be determined from the intercept of that relation. The overall result of that analysis indicated a primordial mass fraction, $Y_p = 0.232 \pm 0.003$. In [43], the stability of this fit was verified by a Monte-Carlo analysis showing that the fits were not overly sensitive to any particular HII region. In addition, the data from [41] was also included, yielding a ^{4}He mass fraction [43]

$$Y_p = 0.234 \pm 0.003 \pm 0.005 \qquad (14)$$

The second uncertainty is an estimate of the systematic uncertainty in the abundance determination.

The ^{7}Li abundance is also reasonably well known. In old, hot, population-II stars, ^{7}Li is found to have a very nearly uniform abundance. For stars with a surface temperature $T >$ 5500 K and a metallicity less than about 1/20th solar (so that effects such as stellar convection may not be important), the abundances show little or no dispersion beyond that which is consistent with the errors of individual measurements. Indeed, as detailed in [44], much of the work concerning ^{7}Li has to do with the presence or absence of dispersion and whether or not there is in fact some tiny slope to a [Li] = log ^{7}Li/H + 12 vs. T or [Li] vs. [Fe/H] relationship. There is ^{7}Li data from nearly 100 halo stars, from a variety of sources. I will use the value given in [45] as the best estimate for the mean ^{7}Li abundance and its statistical uncertainty in halo stars

$$\mathrm{Li/H} = (1.6 \pm 0.1^{+0.4+1.6}_{-0.3-0.5}) \times 10^{-10} \tag{15}$$

The first error is statistical, and the second is a systematic uncertainty that covers the range of abundances derived by various methods. The third set of errors in Eq. (15) accounts for the possibility that as much as half of the primordial ^{7}Li has been destroyed in stars, and that as much as 30% of the observed ^{7}Li may have been produced in cosmic ray collisions rather than in the Big Bang. Observations of ^{6}Li, Be, and B help constrain the degree to which these effects play a role [46]. For ^{7}Li, the uncertainties are clearly dominated by systematic effects.

Turning to D/H, we have three basic types of abundance information: 1) ISM data, 2) solar system information, and perhaps 3) a primordial abundance from quasar absorption systems. The best measurement for ISM D/H is [47]

$$(\mathrm{D/H})_{\mathrm{ISM}} = 1.60 \pm 0.09^{+0.05}_{-0.10} \times 10^{-5} \tag{16}$$

This value may not be universal (or galactic as the case may be) since there may be some real dispersion of D/H in the ISM [48]. The solar abundance of D/H is inferred from two distinct measurements of ^{3}He. The solar wind measurements of ^{3}He as well as the low temperature components of step-wise heating measurements of ^{3}He in meteorites yield the presolar (D + ^{3}He)/H ratio, as D was efficiently burned to ^{3}He in the Sun's pre-main-sequence phase. These measurements indicate that [49, 50]

$$\left(\frac{\mathrm{D} + {}^3\mathrm{He}}{\mathrm{H}}\right)_\odot = (4.1 \pm 0.6 \pm 1.4) \times 10^{-5} \tag{17}$$

The high temperature components in meteorites are believed to yield the true solar ^{3}He/H ratio of [49, 50]

$$\left(\frac{{}^3\mathrm{He}}{\mathrm{H}}\right)_\odot = (1.5 \pm 0.2 \pm 0.3) \times 10^{-5} \tag{18}$$

The difference between these two abundances reveals the presolar D/H ratio, giving,

$$(\mathrm{D/H})_\odot \approx (2.6 \pm 0.6 \pm 1.4) \times 10^{-5} \tag{19}$$

Finally, there have been several recent reported measurements of D/H is high redshift quasar absorption systems. Such measurements are in principle capable of determining the primordial value for D/H and hence η, because of the strong and monotonic dependence of D/H on η. However, at present, detections of D/H using quasar absorption systems indicate both a high

and low value of D/H. As such, we caution that these values may not turn out to represent the true primordial value. The first of these measurements [51] indicated a rather high D/H ratio, D/H $\approx 1.9 - 2.5 \times 10^{-4}$. A recent re-observation of the high D/H absorption system has been resolved into two components, both yielding high values with an average value of D/H = $1.9^{+0.4}_{-0.3} \times 10^{-4}$ [52] as well as an additional system with a similar high value [53]. Other high D/H ratios were reported in [54]. However, there are reported low values of D/H in other such systems [55] with values D/H $\simeq 2.5 \times 10^{-5}$, significantly lower than the ones quoted above. It is probably premature to use this value as the primordial D/H abundance in an analysis of big bang nucleosynthesis, but it is certainly encouraging that future observations may soon yield a firm value for D/H. It is however important to note that there does seem to be a trend that over the history of the Galaxy, the D/H ratio is decreasing, something we expect from galactic chemical evolution. Of course the total amount of deuterium astration that has occurred is still uncertain, and model dependent.

There are also several types of ^{3}He measurements. As noted above, meteoritic extractions yield a presolar value for ^{3}He/H as given in Eq. (18). In addition, there are several ISM measurements of ^{3}He in galactic HII regions [56] which also show a wide dispersion

$$\left(\frac{^3\mathrm{He}}{\mathrm{H}}\right)_{\mathrm{HII}} \simeq 1 - 5 \times 10^{-5} \tag{20}$$

There is also a recent ISM measurement of ^{3}He [57] with

$$\left(\frac{^3\mathrm{He}}{\mathrm{H}}\right)_{\mathrm{ISM}} = 2.1^{+.9}_{-.8} \times 10^{-5} \tag{21}$$

Finally there are observations of ^{3}He in planetary nebulae [58] which show a very high ^{3}He abundance of ^{3}He/H $\sim 10^{-3}$.

Each of the light element isotopes can be made consistent with theory for a specific range in η. Overall consistency of course requires that the range in η agree among all four light elements. ^{3}He (together with D) has stood out in its importance for BBN, because it provided a (relatively large) lower limit for the baryon-to-photon ratio [59], $\eta_{10} > 2.8$. This limit for a long time was seen to be essential because it provided the only means for bounding η from below and in effect allows one to set an upper limit on the number of neutrino flavors [60], N_ν, as well as other constraints on particle physics properties. That is, the upper bound to N_ν is strongly dependent on the lower bound to η. This is easy to see: for lower η, the ^{4}He abundance drops, allowing for a larger N_ν, which would raise the ^{4}He abundance. However, for $\eta < 4 \times 10^{-11}$, corresponding to $\Omega h^2 \sim .001 - .002$, which is not too different from galactic mass densities, there is no bound whatsoever on N_ν [61]. Of course, with the improved data on ^{7}Li, we do have lower bounds on η which exceed 10^{-10}.

In [59], it was argued that since stars (even massive stars) do not destroy ^{3}He in its entirety, we can obtain a bound on η from an upper bound to the solar D and ^{3}He abundances. One can in fact limit [59, 62] the sum of primordial D and ^{3}He by applying the expression below at $t = \odot$

$$\left(\frac{\mathrm{D} + ^3\mathrm{He}}{\mathrm{H}}\right)_{\mathrm{p}} \leq \left(\frac{\mathrm{D}}{\mathrm{H}}\right)_{\mathrm{t}} + \frac{1}{g_3}\left(\frac{^3\mathrm{He}}{\mathrm{H}}\right)_{\mathrm{t}} \tag{22}$$

In (22), g_3 is the fraction of a star's initial D and ^{3}He which survives as ^{3}He. For $g_3 > 0.25$ as suggested by stellar models, and using the solar data on D/H and ^{3}He/H, one finds $\eta_{10} > 2.8$.

The limit $\eta_{10} > 2.8$ derived using (22) is really a one shot approximation. Namely, it is assumed that material passes through a star no more than once. To determine whether or not the solar (and present) values of D/H and ^{3}He/H can be matched it is necessary to consider models of galactic chemical evolution [63]. In the absence of stellar ^{3}He production, particularly by low mass stars, it was shown [64] that there are indeed suitable choices for a star formation rate and an initial mass function to: 1) match the D/H evolution from a primordial value $(D/H)_p = 7.5 \times 10^{-5}$, corresponding to $\eta_{10} = 3$, through the solar and ISM abundances, while 2) at the same time keeping the ^{3}He/H evolution relatively flat so as not to overproduce ^{3}He at the solar and present epochs. This was achieved for $g_3 > 0.3$. Even for $g_3 \sim 0.7$, the present ^{3}He/H could be matched, though the solar value was found to be a factor of 2 too high. For $(D/H)_p \simeq 2 \times 10^{-4}$, corresponding to $\eta_{10} \simeq 1.7$, models could be found which destroy D sufficiently; however, overproduction of ^{3}He occurred unless g_3 was tuned down to about 0.1.

In the context of models of galactic chemical evolution, there is, however, little justification a priori for neglecting the production of ^{3}He in low mass stars. Indeed, stellar models predict that considerable amounts of ^{3}He are produced in stars between 1 and 3 $M_\odot$. For $M < 8M_\odot$, Iben and Truran [65] calculate

$$(^3\mathrm{He/H})_f = 1.8 \times 10^{-4} \left(\frac{M_\odot}{M}\right)^2 + 0.7 \left[(\mathrm{D} + {}^3\mathrm{He})/\mathrm{H}\right]_i \tag{23}$$

so that at $\eta_{10} = 3$, and $((D + {}^3He)/H)_i = 9 \times 10^{-5}$, $g_3(1M_\odot) = 2.7$! It should be emphasized that this prediction is in fact consistent with the observation of high ^{3}He/H in planetary nebulae [58].

Generally, implementation of the ^{3}He yield in Eq. (23) in chemical evolution models leads to an overproduction of ^{3}He/H particularly at the solar epoch [66, 67]. In Figure 1, the evolution of D/H and ^{3}He/H is shown as a function of time from [49, 66]. The solid curves show the evolution in a simple model of galactic chemical evolution with a star formation rate proportional to the gas density and a power law IMF (see [66]) for details). The model was chosen to fit the observed deuterium abundances. However, as one can plainly see, ^{3}He is grossly overproduced (the deuterium data is represented by squares and ^{3}He by circles). Depending on the particular model chosen, it may be possible to come close to at least the upper end of the range of the ^{3}He/H observed in galactic HII regions [56], although the solar value is missed by many standard deviations.

The overproduction of ^{3}He relative to the solar meteoritic value seems to be a generic feature of chemical evolution models when ^{3}He production in low mass stars is included. In [49], a more extreme model of galactic chemical evolution was tested. There, it was assumed that the initial mass function was time dependent in such a way so as to favor massive stars early on (during the first two Gyr of the galaxy). Massive stars are preferential from the point of view of destroying ^{3}He. However, massive stars are also proficient producers of heavy elements and in order to keep the metallicity of the disk down to acceptable levels, supernovae driven outflow was also included. The degree of outflow was limited roughly by the observed metallicity in the intergalactic gas in clusters of galaxies. One further assumption was necessary; we allowed the massive stars to lose their ^{3}He depleted hydrogen envelopes prior to explosion. Thus only the heavier elements were expulsed from the galaxy. With all of these (semi-defensible) assumptions, ^{3}He was still overproduced at the solar epoch, as shown by the dashed curve in Figure 1. Though there certainly is an improvement in the evolution of ^{3}He without reducing the yields of low mass stars, it is hard to envision much further reduction in the solar ^{3}He predicted by these

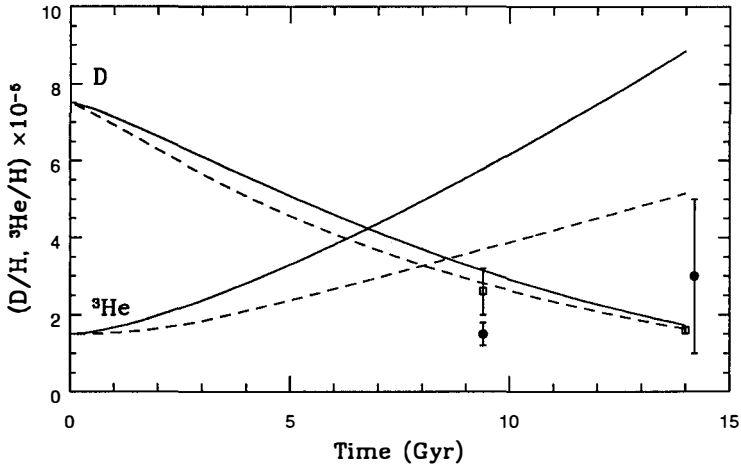


Figure 1: The evolution of D and ^{3}He with time.

models. The only conclusion that we can make at this point is that there is clearly something wrong with our understanding of ^{3}He in terms of either chemical evolution, stellar evolution or perhaps even the observational data.

Given the magnitude of the problems concerning ^{3}He, it would seem unwise to make any strong conclusion regarding big bang nucleosynthesis which is based on ^{3}He. Perhaps as well some caution is deserved with regard to the recent D/H measurements, although if the present trend continues and is verified in several different quasar absorption systems, then D/H will certainly become our best measure for the baryon-to-photon ratio. Given the current situation however, it makes sense to take a step back and perform an analysis of big bang nucleosynthesis in terms of the element isotopes that are best understood, namely ^{4}He and ^{7}Li.

Monte Carlo techniques are proving to be a useful form of analysis for big bang nucleosynthesis [68, 69]. In [18], we performed just such an analysis using only ^{4}He and ^{7}Li. It should be noted that in principle, two elements should be sufficient for constraining the one parameter (η) theory of BBN. We begin by establishing likelihood functions for the theory and observations. For example, for ^{4}He, the theoretical likelihood function takes the form

$$L_{\rm BBN}(Y, Y_{\rm BBN}) = e^{-(Y - Y_{\rm BBN}(\eta))^2/2\sigma_1^2} \tag{24}$$

where $Y_{\rm BBN}(\eta)$ is the central value for the ^{4}He mass fraction produced in the big bang as predicted by the theory at a given value of η, and σ_1 is the uncertainty in that value derived from the Monte Carlo calculations [69] and is a measure of the theoretical uncertainty in the big bang calculation. Similarly one can write down an expression for the observational likelihood function. In this case we have two sources of errors, as discussed above, a statistical uncertainty, σ_2 and a systematic uncertainty, $\sigma_{\rm sys}$. Here, I will assume that the systematic error is described by a top hat distribution [69, 70]. The convolution of the top hat distribution and the Gaussian (to describe the statistical errors in the observations) results in the difference of two error

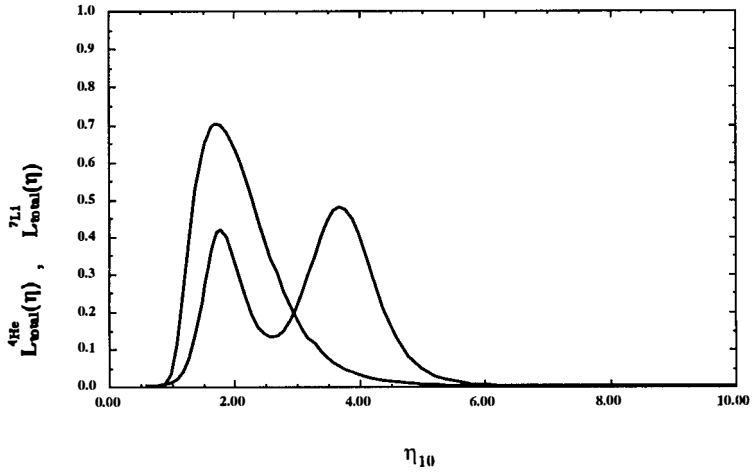


Figure 2: Likelihood distribution for each of ^{4}He and ^{7}Li, shown as a function of η. The one-peak structure of the ^{4}He curve corresponds to its monotonic increase with η, while the two-peaks for ^{7}Li arise from its passing through a minimum.

functions

$$L_{\rm O}(Y, Y_{\rm O}) = \mathrm{erf}\left(\frac{Y - Y_{\rm O} + \sigma_{\rm sys}}{\sqrt{2}\sigma_2}\right) - \mathrm{erf}\left(\frac{Y - Y_{\rm O} - \sigma_{\rm sys}}{\sqrt{2}\sigma_2}\right) \tag{25}$$

where in this case, $Y_{\rm O}$ is the observed (or observationally determined) value for the ^{4}He mass fraction. (Had I used a Gaussian to describe the systematic uncertainty, the convolution of two Gaussians leads to a Gaussian, and the likelihood function (25) would have taken a form similar to that in (24).

A total likelihood function for each value of η_{10} is derived by convolving the theoretical and observational distributions, which for ^{4}He is given by

$$L^{^4\mathrm{He}}{}_{\rm total}(\eta) = \int dY\, L_{\rm BBN}\left(Y, Y_{\rm BBN}\left(\eta\right)\right) L_{\rm O}(Y, Y_{\rm O}) \tag{26}$$

An analogous calculation is performed [18] for ^{7}Li. The resulting likelihood functions from the observed abundances given in Eqs. (14) and (15) is shown in Figure 2. As one can see there is very good agreement between ^{4}He and ^{7}Li in the vicinity of $\eta_{10} \simeq 1.8$.

The combined likelihood, for fitting both elements simultaneously, is given by the product of the two functions in Figure 2 and is shown in Figure 3. From Figure 2 it is clear that ^{4}He overlaps the lower (in η) ^{7}Li peak, and so one expects that there will be concordance in an allowed range of η given by the overlap region. This is what one finds in Figure 3, which does show concordance and gives a preferred value for η, $\eta_{10} = 1.8^{+1}_{-.2}$ corresponding to $\Omega_B h^2 = .006^{+.004}_{-.001}$.

Thus, we can conclude that the abundances of ^{4}He and ^{7}Li are consistent, and select an η_{10} range which overlaps with (at the 95% CL) the longstanding favorite range around $\eta_{10} = 3$. Furthermore, by finding concordance using only ^{4}He and ^{7}Li, we deduce that if there is problem with BBN, it must arise from D and ^{3}He and is thus tied to chemical evolution or the

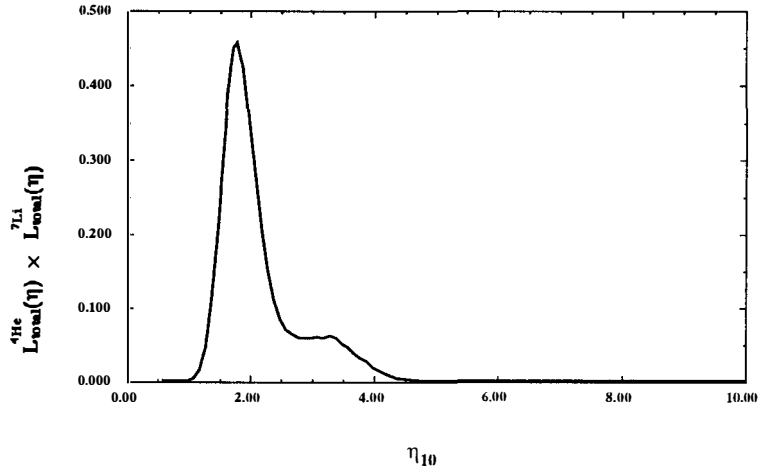


Figure 3: Combined likelihood for simultaneously fitting ^{4}He and ^{7}Li, as a function of η.

stellar evolution of ^{3}He. The most model-independent conclusion is that standard BBN with $N_\nu = 3$ is not in jeopardy, but there may be problems with our detailed understanding of D and particularly ^{3}He chemical evolution. It is interesting to note that the central (and strongly) peaked value of η_{10} determined from the combined ^{4}He and^{7}Li likelihoods is at $\eta_{10} = 1.8$. The corresponding value of D/H is 1.8×10^{-4}, very close to the high value of D/H in quasar absorbers [51, 52, 54]. Since D and ^{3}He are monotonic functions of η, a prediction for η, based on ^{4}He and ^{7}Li, can be turned into a prediction for D and ^{3}He. The corresponding 95% CL ranges are D/H $= (5.5 - 27) \times 10^{-5}$ and and ^{3}He/H $= (1.4 - 2.7) \times 10^{-5}$.

If we did have full confidence in the measured value of D/H in quasar absorption systems, then we could perform the same statistical analysis using ^{4}He, ^{7}Li, and D. To include D/H, one would proceed in much the same way as with the other two light elements. We compute likelihood functions for the BBN predictions as in Eq. (24) and the likelihood function for the observations using D/H $= (1.9 \pm 0.4) \times 10^{-4}$. We are using only the high value of D/H here. These are then convolved as in Eq. (26). In figure 4, the resulting normalized distribution, $L^{\rm D}_{\rm total}(\eta)$ is super-imposed on distributions appearing in figure 2. It is indeed startling how the three peaks, for D, ^{4}He and ^{7}Li are literally on top of each other. In figure 5, the combined distribution is shown. We now have a very clean distribution and prediction for η, $\eta_{10} = 1.75^{+.3}_{-.1}$ corresponding to $\Omega_B h^2 = .006^{+.001}_{-.0004}$, with the peak of the distribution at $\eta_{10} = 1.75$. The absence of any overlap with the high-η peak of the ^{7}Li distribution has considerably lowered the upper limit to η. Overall, the concordance limits in this case are dominated by the deuterium likelihood function.

To summarize on the subject of big bang nucleosynthesis, I would assert that one can conclude that the present data on the abundances of the light element isotopes are consistent with the standard model of big bang nucleosynthesis. Using the the isotopes with the best data, ^{4}He and ^{7}Li, it is possible to constrain the theory and obtain a best value for the baryon-to-

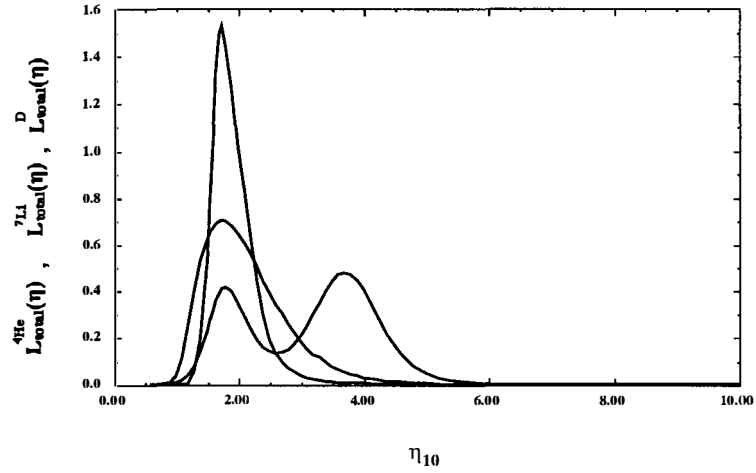


Figure 4: As in Figure 2, with the addition of the likelihood distribution for D/H.

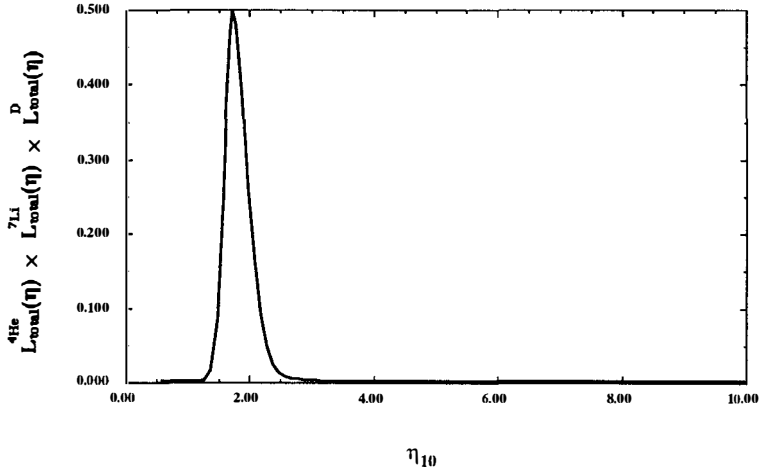


Figure 5: Combined likelihood for simultaneously fitting ^{4}He and ^{7}Li, and D as a function of η.

photon ratio of $\eta_{10} = 1.8$, a corresponding value $\Omega_B = .0065$ and

$$\begin{aligned} 1.4 &< \eta_{10} &< 3.8 \quad 95\%\text{CL} \\ .005 &< \Omega_B h^2 &< .014 \quad 95\%\text{CL} \end{aligned} \tag{27}$$

For $0.4 < h < 1$, we have a range $.005 < \Omega_B < .09$. This is a rather low value for the baryon density and would suggest that much of the galactic dark matter is non-baryonic [71]. These predictions are in addition consistent with recent observations of D/H in quasar absorption systems which show a high value. Difficulty remains however, in matching the solar ^{3}He abundance, suggesting a problem with our current understanding of galactic chemical evolution or the stellar evolution of low mass stars as they pertain to ^{3}He.

If we now take as given that non-baryonic dark matter is required, we are faced with the problem of its identity. Light neutrinos ($m \leq 30eV$) are a long-time standard when it comes to non-baryonic dark matter [72]. Light neutrinos produce structure on large scales and the natural (minimal) scale for structure clustering is given in Eq. (11). Hence neutrinos offer the natural possibility for large scale structures [73, 74] including filaments and voids. It seemed, however, that neutrinos were ruled out because they tend to produce too much large scale structure [75]. Because the smallest non-linear structures have mass scale M_J and the typical galactic mass scale is $\simeq 10^{12} M_\odot$, galaxies must fragment out of the larger pancake-like objects. The problem is that in such a scenario, galaxies form late [74, 76] ($z \leq 1$) whereas quasars and galaxies are seen out to redshifts $z \gtrsim 4$. Recently, neutrinos are seeing somewhat of a revival in popularity in mixed dark matter models.

In the standard model, the absence of a right-handed neutrino state precludes the existence of a neutrino mass. By adding a right-handed state ν_R, it is possible to generate a Dirac mass for the neutrino, $m_\nu = h_\nu v/\sqrt{2}$, as for the charged lepton masses, where h_ν is the neutrino Yukawa coupling constant, and v is the Higgs expectation value. It is also possible to generate a Majorana mass for the neutrino when in addition to the Dirac mass term, $m_\nu \bar{\nu_R} \nu_L$, a term $M \nu_R \nu_R$ is included. In what is known as the see-saw mechanism, the two mass eigenstates are given by $m_{\nu_1} \sim m_\nu^2/M$ which is very light, and $m_{\nu_2} \sim M$ which is heavy. The state ν_1 is our hot dark matter candidate as ν_2 is in general not stable.

The cosmological constraint on the mass of a light neutrino is derived from the overall mass density of the Universe. In general, the mass density of a light particle χ can be expressed as

$$\rho_\chi = m_\chi Y_\chi n_\gamma \leq \rho_c = 1.06 \times 10^{-5} h_o{}^2 \text{GeV/cm}^3 \tag{28}$$

where $Y_\chi = n_\chi/n_\gamma$ is the density of χ's relative to the density of photons, for $\Omega h_o{}^2 < 1$. For neutrinos $Y_\nu = 3/11$, and one finds [77]

$$\sum_\nu (\frac{g_\nu}{2}) m_\nu < 93\text{eV}(\Omega h_o{}^2) \tag{29}$$

where the sum runs over neutrino flavors. All particles with abundances Y similar to neutrinos will have a mass limit given in Eq. (29).

It was possible that neutrinos (though not any of the known flavors) could have had large masses, $m_\nu > 1$ MeV. In that case their abundance Y is controlled by $\nu, \bar{\nu}$ annihilations [78], for example, $\nu\bar{\nu} \to f\bar{f}$ via Z exchange. When the annihilations freeze-out (the annihilation rate becomes slower than the expansion rate of the Universe), Y becomes fixed. Roughly, $Y \sim (m\sigma_A)^{-1}$ and $\rho \sim \sigma_A{}^{-1}$ where σ_A is the annihilation cross-section. For neutrinos, we

expect $\sigma_A \sim m_\nu{}^2/m_Z{}^4$ so that $\rho_\nu \sim 1/m_\nu{}^2$ and we can derive a lower bound [79, 80] on the neutrino mass, $m_\nu \gtrsim 3-7$ GeV, depending on whether it is a Dirac or Majorana neutrino. Indeed, any particle with roughly a weak scale cross-sections will tend to give an interesting value of $\Omega h^2 \sim 1$ [81].

Due primarily to the limits from LEP [82], the heavy massive neutrino has become simply an example and is no longer a dark matter candidate. LEP excludes neutrinos (with standard weak interactions) with masses $m_\nu \lesssim 40$ GeV. Lab constraints for Dirac neutrinos are available [83], excluding neutrinos with masses between 10 GeV and 4.7 TeV. This is significant, since it precludes the possibility of neutrino dark matter based on an asymmetry between ν and $\bar{\nu}$ [84]. Majorana neutrinos are excluded as *dark matter* since $\Omega_\nu h_o{}^2 < 0.001$ for $m_\nu > 40$ GeV and are thus cosmologically uninteresting.

Supersymmetric theories introduce several possible candidates. If R-parity, which distinguishes between "normal" matter and the supersymmetric partners and can be defined in terms of baryon, lepton and spin as $R = (-1)^{3B+L+2S}$, is unbroken, there is at least one supersymmetric particle which must be stable. I will assume R-parity conservation, which is common in the MSSM. R-parity is generally assumed in order to justify the absence of superpotential terms can be responsible for rampid proton decay. The stable particle (usually called the LSP) is most probably some linear combination of the only $R = -1$ neutral fermions, the neutralinos [85]: the wino $\tilde{W}^3$, the partner of the 3rd component of the $SU(2)_L$ gauge boson; the bino, $\tilde{B}$, the partner of the $U(1)_Y$ gauge boson; and the two neutral Higgsinos, $\tilde{H}_1$ and $\tilde{H}_2$. Gluinos are expected to be heavier -$m_{\tilde{g}} = (\frac{\alpha_3}{\alpha}) \sin^2 \theta_W M_2$, where M_2 is the supersymmetry breaking SU(2) gaugino mass- and they do not mix with the other states. The sneutrino [86] is also a possibility but has been excluded as a dark matter candidate by direct [83] searches, indirect [87] and accelerator[82] searches. For more on the motivations for supersymmetry and the supersymmetric parameter space, see the contribution of Jungman [81].

The identity of the LSP is effectively determined by three parameter in the MSSM, the gaugino mass, M_2, the Higgs mixing mass μ, and the ratio of the Higgs vacuum expectation values, $\tan\beta$. In Figure 6 [89], regions in the M_2, μ plane with $\tan\beta = 2$ are shown in which the LSP is one of several nearly pure states, the photino, $\tilde{\gamma}$, the U(1) gaugino, $\tilde{B}$, a symmetric combination of the Higgsinos, $\tilde{H}_{(12)} = \frac{1}{\sqrt{2}}(\tilde{H}_1 + \tilde{H}_2)$, or the Higgsino $\tilde{S} = \tilde{H}_1 \cos\beta + \tilde{H}_2 \sin\beta$. The dashed lines show the LSP mass contours. The cross hatched regions correspond to parameters giving a chargino ($\tilde{W}^\pm$, $\tilde{H}^\pm$) state with mass $m_{\tilde{\chi}} \leq 45 GeV$ and as such are excluded by LEP[90]. This constraint has been extended by LEP1.5, [91, 92] and is shown by the light shaded region and corresponds to regions where the chargino mass is $\lesssim 67$ GeV. The dark shaded region corresponds to a limit on M_2 from the limit[93] on the gluino mass $m_{\tilde{g}} \leq 70$ GeV or $M_2 \leq 22$ GeV. Notice that the parameter space is dominated by the $\tilde{B}$ or $\tilde{H}_{12}$ pure states and that the photino (most often quoted as the LSP) only occupies a small fraction of the parameter space, as does the Higgsino combination $\tilde{S}^0$.

As described in [81], the relic abundance of LSP's is determined by solving the Boltzmann equation for the LSP number density in an expanding Universe. The technique[80] used is similar to that for computing the relic abundance of massive neutrinos[78]. For binos, as was the case for photinos [88], it is possible to adjust the sfermion masses $m_{\tilde{f}}$ to obtain closure density. Adjusting the sfermion mixing parameters [95] or CP violating phases [96] allows even greater freedom. In Figure 7 [97], the relic abundance (Ωh^2) is shown in the $M_2 - \mu$ plane with $\tan\beta = 2$, the Higgs pseudoscalar mass $m_0 = 50$ GeV, $m_t = 100$ GeV, and $m_{\tilde{f}} = 200$ GeV. Clearly the MSSM offers sufficient room to solve the dark matter problem. Similar results have

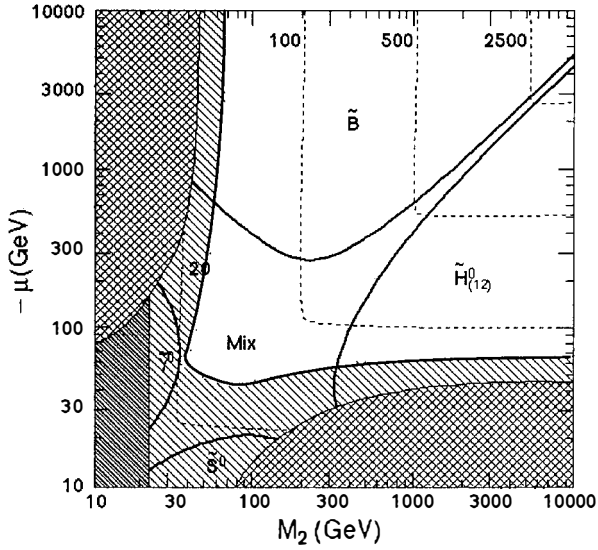


Figure 6: The M_2-μ plane in the MSSM for $\tan\beta = 2$.

been found by other groups [98, 99, 100]. In Figure 7, in the higgsino sector $\tilde{H}_{12}$ marked off by the dashed line, co-annihilations [101, 99] between $\tilde{H}_{(12)}$ and the next lightest neutralino (also a Higgsino) were not included. These tend to lower significantly the relic abundance in much of this sector.

Though I have concentrated on the LSP in the MSSM as a cold dark matter candidate, there are many other possibilities when one goes beyond the standard model. Axions were discussed at length by Jungman [81] and Lu [102]. A host of other possibilities were discussed by Khlopov [103].

The final subject that I will cover in this introduction/summary is the question of detection. Dark matter detection can be separated into two basic methods, direct [104] and indirect [105]. Direct detection relies on the ability to detect the elastic scattering of a dark matter candidate off a nucleus in a detector. The experimental signatures for direct detection were covered by Cabrera [106] and several individual experiments were described [107].

The detection rate, will depend on the density of dark matter in the solar neighborhood, $\rho \sim 0.3$ GeV/cm^3, the velocity, $v \sim 300$ km/s, and the elastic cross section, σ. Spin independent interactions are the most promising for detection. Dirac neutrinos have spin-independent interactions, but as noted above, these have already been excluded as dark matter by direct detection experiments [83]. In the MSSM, it is possible for the LSP to also have spin independent interactions which are mediated by Higgs exchange. These scatterings are only important when the LSP is a mixed (gaugino/Higgsino) state as in the central regions of Figures 6 and 7. Generally, these regions have low values of Ωh^2 (since the annihilation cross sections are also enhanced) and the parameter space in which the elastic cross section and relic density are large

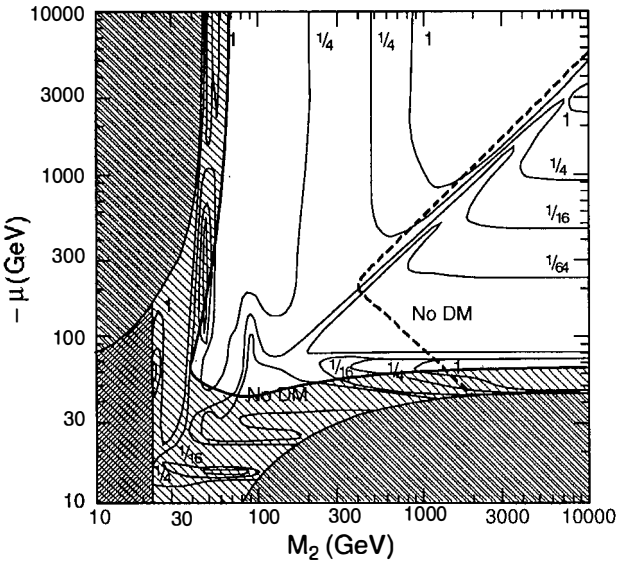


Figure 7: Relic density contours (Ωh^2) in the M_2 - μ plane.

is rather limited. Furthermore, a significant detection rate in this case relies on a low mass for the Higgs scalar [108, 109].

More typical of the SUSY parameter space is a LSP with spin dependent interactions. Elastic scatterings are primarily spin dependent whenever the LSP is mostly either gaugino or Higgsino. For Higgsino dark matter, Higgsinos with scatterings mediated by Z^0 avoid the $\tilde{H}_{(12)}$ regions of Figures 6 and 7, and as such are now largely excluded (the S^0 region does grow at low $\tan\beta$ [85, 89]. Higgsino scatterings mediated by sfermion exchange depend on couplings proportional to the light quark masses and will have cross sections which are suppressed by $(m_p/m_W)^4$, where m_p is the proton mass. These rates are generally very low [109]. Binos, on the other hand, will have elastic cross sections which go as $m^2/m_{\tilde{f}}^4$, where m is the reduced mass of the bino and nucleus. These rates are typically higher (reaching up to almost 0.1 events per kg-day [109, 110, 111].

Indirect methods also offer the possibility for the detection of dark matter. Three methods for indirect detection were discussed [105]. 1) γ-rays from dark matter annihilations in the galactic halo are a possible signature [112]. In the case of the MSSM, unless the mass of the LSP is larger than m_W, the rates are probably too small to be detectable over background [105]. 2) Dark matter will be trapped gradually in the sun, and annihilations within the sun will produce high energy neutrinos which may be detected [113]; similarly, annihilations within the earth may provide a detectable neutrino signal [114]. Edsjo [115] discussed possibilities for determining the mass the dark matter candidate from the angular distribution of neutrinos. This method hold considerable promise, as there will be a number of very large neutrino detectors coming on line in the future. Finally, 3) there is the possibility that halo annihilations into positrons

and antiprotons in sufficient numbers to distinguish them from cosmic-ray backgrounds [112, 116, 117].

Acknowledgments

I would like to thank T. Falk for his help in proof reading this manuscript. This work was supported in part by DOE grant DE–FG02–94ER–40823.

References

[1] see: J.R.. Primack in *Enrico Fermi. Course 92*, ed. N. Cabibbo (North Holland, Amsterdam, 1987), p. 137; V. Trimble, *Ann. Rev. Astron. Astrophys.* **25** (1987) 425; J. Primack, D. Seckel, and B. Sadulet, *Ann. Rev. Nucl. Part. Sci.* **38** (1988) 751; *Dark Matter*, ed. M. Srednicki (North-Holland, Amsterdam,1989).

[2] J.L.Tonry, in *Relativistic Astrophysics and Particle Cosmology* ed. C.W. Akerlof and M. Srednicki (New York Academy of Sciences, New York, 1993), p.113.

[3] S.M. Faber and J.J. Gallagher, *Ann. Rev. Astron. Astrophys.* **17** (1979) 135.

[4] A. Bosma,*Ap.J.* **86** (1981) 1825.

[5] V.C. Rubin, W.K. Ford and N. Thonnard, *Ap.J.* **238** (1980) 471; V.C. Rubin, D. Burstein, W.K. Ford and N. Thonnard, *Ap.J.* **289** (1985) 81; T.S. Van Albada and R.. Sancisi, *Phil. Trans. R. Soc. Land.* **A320** (1986) 447.

[6] M. Persic and P. Salucci, *ApJ Supp* **99** (1995) 501; M. Persic, P. Salucci, and F. Stel, to be published in *Dark Matter* Proc. of the 5th Annual October Conference in Maryland (1995).

[7] R.P. Saglia et al., *Ap.J.* **403** (1993) 567; C.M. Carollo et al., *Ap.J.* **411** (1995) 25;

[8] D. Fabricant and P. Gorenstein, *Ap.J.* **267** (1983) 535; G.C. Stewart, C.R. Canizares, A.C. Fabian and P.E.J. Nilsen, *Ap.J.* **278** (1984) 53 and references therein.

[9] W. Foreman, these proceedings.

[10] R. Mushotzky, in *Relativistic Astrophysics and Particle Cosmology* ed. C.W. Akerlof and M. Srednicki (New York Academy of Sciences, New York, 1993), p. 184; J.S. Mulchaey, D.S. Davis, R.F. Mushotzky, and D. Burstein, *ApJ* **404** (1993) L9; M.J. Henriksen and G.A. Mamon, *Ap.J.* **421** (1994) L63.

[11] L.P David, C. Jones, and W. Foreman, *Ap.J.* **445** (1995) 578.

[12] P.J.E. Peebles, *The Large Scale Structure of the Universe*, (Princeton University Press, Princeton, 1980).

[13] A. Dressler, D. Lynden-Bell D. Burstein, R. Davies, S. Faber, R. Terlevich, and G. Wegner, *Ap.J.* **313** (1987) L37; A. Deckel, E. Bertschinger, A. Yahil, M.A. Strauss, M. Davis, and J.P. Huchra, *Ap.J.* **412** (1993) 1.

[14] J.A. Tyson, F. Valdes, and R.A. Wenk, *Ap.J.* **349** (1990) L1.

[15] L. Van Waerbeke, these proceedings.

[16] for a review see: K.A. Olive, *Phys. Rep.* **190** (1990) 181.

[17] T.P. Walker, G. Steigman, D.N. Schramm, K.A. Olive and K. Kang, *Ap.J.* **376** (1991) 51.

[18] B.D. Fields and K.A. Olive, *Phys. Lett.* **B368** (1996) 103; B.D. Fields, K. Kainulainen, D. Thomas, and K.A. Olive, astro-ph/9603009, *New Astronomy* (1996) in press.

[19] W.H. Press, *Phys. Scr.* **21** (1980) 702; V.F. Mukhanov and G.V. Chibisov, *JETP Lett.* **33** (1981) 532; S.W. Hawking, *Phys. Lett.* **115B** (1982) 295; A.A. Starobinsky, , *Phys. Lett.* **117B** (1982) 175; A.H. Guth and S.Y. Pi, *Phys. Rev. Lett.* **49** (1982) 1110; J.M. Bardeen, P.J. Steinhardt and M.S. Turner, *Phys. Rev.* **D28** (1983) 679.

[20] E.R. Harrison, *Phys. Rev.* **D1** (1979) 2726; Ya.B. Zeldovich, *M.N.R.A.S.* **160** (1972) 1P.

[21] M. Lachieze-Rey, these proceedings.

[22] F. Bernardeau, these proceedings.

[23] S. Colombi, these proceedings.

[24] G. Tormen, astro-ph/9604081, these proceedings.

[25] G. Kauffmann, these proceedings.

[26] E.L. Wright et al. *Ap.J.* **396** (1992) L13; G. Hinshaw et al., astro-ph/9601058 (1996).

[27] J.R. Bond and A.S. Szalay, *Ap.J.* **274** (1983) 443.

[28] J.R. Bond, G. Efstathiou and J. Silk, *Phys. Lett.* **45** (1980) 1980; Ya.B. Zeldovich and R.A.Sunyaev, *Son. Ast. Lett.* **6** (1980) 457.

[29] M. Vogeley, these proceedings.

[30] J. Loveday, astro-ph/9605028, these proceedings; G. Mamon, these proceedings.

[31] R. Mandelosi, these proceedings.

[32] W. Zurek, these proceedings.

[33] J. Delabrouille, these proceedings.

[34] Hegyi, D.J. and Olive, K.A., *Phys. Lett.***126B** (1983) 28; *Ap. J.* **303** (1986) 56.

[35] B. Paczynski, *Ap.J.* **304** (1986) 1.

[36] C. Alcock et al., *Nature* **365** (1983) 621; E. Aubourg et al. *Nature* **365** (1983) 623.

[37] C. Alcock et al., *Phys. Rev. Lett.* **74** (1995) 2867; astro-ph/9506113 (1995); E. Aubourg et al., *A.A.* **301** (1995) 1.

[38] D.P. Bennett et al., astro-ph/9510104 (1995).

[39] B.E.J. Pagel, E.A. Simonson, R.J. Terlevich and M. Edmunds, *MNRAS* **255** (1992) 325.

[40] E. Skillman et al., *Ap.J. Lett.* (in preparation) 1995.

[41] Y.I. Izatov, T.X. Thuan, and V.A. Lipovetsky, *Ap.J.* **435** (1994) 647.

[42] K.A. Olive and G. Steigman, *Ap.J. Supp.* **97** (1995) 49.

[43] K.A. Olive and S.T. Scully, *IJMPA* 11 (1995) 409.

[44] M. Spite, P. Francois, P.E. Nissen, and F. Spite, *A.A.* **307** (1996) 172; F. Spite, to be published in the Proceedings of the IInd Rencontres du Vietnam: The Sun and Beyond, ed. J. Tran Thanh Van, 1996.

[45] P. Molaro, F. Primas, and P. Bonifacio, *A.A.* **295** (1995) L47.

[46] T.P. Walker, G. Steigman, D.N. Schramm, K.A. Olive and B. Fields, *Ap.J.* **413** (1993) 562; K.A. Olive, and D.N. Schramm, *Nature* **360** (1993) 439; G. Steigman, B. Fields, K.A. Olive, D.N. Schramm, and T.P. Walker, Ap.J. **415** (1993) L35.

[47] J.L. Linsky, et al., *Ap.J.* **402** (1993) 694; J.L. Linsky, et al.,*Ap.J.* **451** (1995) 335.

[48] R. Ferlet, to be published in the Proceedings of the IInd Rencontres du Vietnam: The Sun and Beyond, ed. J. Tran Thanh Van, 1996.

[49] S.T. Scully, M. Cassé, K.A. Olive, D.N. Schramm, J.W. Truran, and E. Vangioni-Flam, astro-ph/0508086, *Ap.J.* **462** (1996) 960.

[50] J. Geiss, in *Origin and Evolution of the Elements*, eds. N. Prantzos, E. Vangioni-Flam, and M. Cassé (Cambridge: Cambridge University Press, 1993), p. 89

[51] R.F. Carswell, M. Rauch, R.J. Weymann, A.J. Cooke, and J.K. Webb, *MNRAS* **268** (1994) L1; A. Songaila, L.L. Cowie, C. Hogan, and M. Rugers, *Nature* **368** (1994) 599.

[52] M. Rugers and C.J. Hogan, *Ap.J.* **459** (1996) L1.

[53] M. Rugers and C.J. Hogan, astro-ph/9603084 (1996).

[54] R.F. Carswell, et al. *MNRAS* **278** (1996) 518; E.J. Wampler, et al. astro-ph/9512084, *A.A.* (1996) in press.

[55] D. Tytler, X.-M. Fan, and S. Burles, astro-ph/9603069 (1996); S. Burles and D. Tytler, astro-ph/9603070 (1996).

[56] D.S. Balser, T.M. Bania, C.J. Brockway, R.T. Rood, and T.L. Wilson, *Ap.J.* **430** (1994) 667.

[57] G. Gloeckler, and J. Geiss, *Nature* (1996) submitted.

[58] R.T. Rood, T.M. Bania, and T.L. Wilson, Nature **355** (1992) 618; R.T. Rood, T.M. Bania, T.L. Wilson, and D.S. Balser, 1995, in *the Light Element Abundances, Proceedings of the ESO/EIPC Workshop*, ed. P. Crane, (Berlin:Springer), p. 201.

[59] J. Yang, M.S. Turner, G. Steigman, D.N. Schramm, and K.A. Olive, *Ap.J.* **281** (1984) 493.

[60] G. Steigman, D.N. Schramm, and J. Gunn, *Phys. Lett.* **B66** (1977) 202.

[61] K.A. Olive, D.N. Schramm, G. Steigman, M.S. Turner, and J. Yang, *Ap.J.* **246** (1981) 557.

[62] D. Black, *Nature Physical Sci.*, **24** (1971) 148.

[63] B.M. Tinsley, *Fund Cosm Phys* **5** (1980) 287.

[64] E. Vangioni-Flam, K.A. Olive, and N. Prantzos, *Ap.J.* **427** (1994) 618.

[65] I. Iben, and J.W. Truran, *Ap.J.* **220** (1978) 980.

[66] K.A. Olive, R.T. Rood, D.N. Schramm, J.W. Truran, and E. Vangioni-Flam, *Ap.J.* **444** (1995) 680.

[67] D. Galli, F. Palla, F. Ferrini, and U. Penco, Ap.J. **443** (1995) 536; D. Dearborn, G. Steigman, and M. Tosi, *Ap.J.* **465** (1996) in press.

[68] L.M. Krauss and P. Romanelli, *Ap.J.* **358** (1990) 47; P.J. Kernan and L.M. Krauss, *Phys. Rev. Lett.* **72** (1994) 3309; L.M. Krauss and P.J. Kernan, *Phys. Lett.* **B347** (1995) 347; M. Smith, L. Kawano, and R.A. Malaney, *Ap.J. Supp.* **85** (1993) 219.

[69] N. Hata, R.J. Scherrer, G. Steigman, D. Thomas, and T.P. Walker, *Ap.J.* **458** (1996) 637.

[70] K.A. Olive and G. Steigman, *Phys. Lett.* **B354** (1995) 357.

[71] E. Vangioni-Flam and M. Cassé, *Ap.J.* **441** (1995) 471.

[72] D.N. Schramm and G. Steigman, *Ap. J.* **243** (1981) 1.

[73] P.J.E. Peebles, *Ap. J.* **258** (1982) 415; A. Melott, *M.N.R.A.S.* **202** (1983) 595; A.A. Klypin, S.F. Shandarin, *M.N.R.A.S.* **204** (1983) 891.

[74] C.S. Frenk, S.D.M. White and M.Davis, *Ap. J.* **271** (1983) 417.

[75] S.D.M. White, C.S. Frenk and M. Davis, *Ap. J.* **274** (1983) 61.

[76] J.R. Bond, J. Centrella, A.S. Szalay and J. Wilson, in *Formation and Evolution of Galaxies and Large Structures in the Universe*, ed. J. Andouze and J. Tran Thanh Van, (Dordrecht-Reidel 1983) p. 87.

[77] R. Cowsik and J. McClelland, *Phys. Rev. Lett.* **29** (1972) 669; A.S. Szalay, G. Marx, *Astron. Astrophys.* **49** (1976) 437.

[78] P. Hut, *Phys. Lett.* **69B** (1977) 85; B.W. Lee and S. Weinberg, *Phys. Rev. Lett.* **39** (1977) 165.

[79] E.W. Kolb and K.A. Olive, *Phys. Rev.* **D33** (1986) 1202; E: **34** (1986) 2531; L.M. Krauss, *Phys. Lett.* **128B** (1983) 37.

[80] R. Watkins, M. Srednicki and K.A. Olive, *Nucl. Phys.* **B310** (1988) 693.

[81] G. Jungman, these proceedings.

[82] see e.g. B. Adeva, et. al., *Phys. Lett.* **B231** (1989) 509; D. Decamp, et. al., *Phys. Lett.* **B231** (1989) 519; M.Z. Akrawy, et. al., *Phys. Lett.* **B231** (1989) 530; P. Aarnio, *Phys. Lett.* **B231** (1989) 539.

[83] S. Ahlen, et. al., *Phys. Lett.* **B195** (1987) 603; D.D. Caldwell, et. al., *Phys. Rev. Lett.* **61** (1988) 510; M. Beck et al., *Phys. Lett.* **B336** (1994) 141.

[84] P. Hut and K.A. Olive, *Phys. Lett.* **B87** (1979) 144.

[85] J.Ellis, J. Hagelin, D.V. Nanopoulos, K.A.Olive and M. Srednicki, *Nucl. Phys.* **B238** (1984) 453.

[86] L.E. Ibanez, *Phys. Lett.* **137B** (1984) 160; J. Hagelin, G.L. Kane, and S. Raby, *Nucl., Phys.* **B241** (1984) 638; T. Falk, K.A. Olive, and M. Srednicki, *Phys. Lett.* **B339** (1994) 248.

[87] see e.g. K.A. Olive and M. Srednicki, *Phys. Lett.* **205B** (1988) 553.

[88] H. Golberg, *Phys. Rev. Lett.* **50** (1983) 1419; L.M. Krauss, *Nucl. Phys.* bf B227 (1983) 556.

[89] K.A. Olive and M. Srednicki, *Phys. Lett.* **B230** (1989) 78; *Nucl. Phys.* **B355** (1991) 208.

[90] ALEPH collaboration,D. Decamp et al., *Phys. Rep.* **216** (1992) 253.

[91] ALEPH collaboration, D. Buskulic et al., CERN-PPE/96-10 (1996).

[92] J. Ellis, T. Falk, K.A. Olive, M. Schmitt, in preparation (1996).

[93] J. Alitti et al., *Phys. Lett.* **B235** (1990) 363.

[94] J. Ellis, G. Ridolfi and F. Zwirner, *Phys. Lett.* **B237** (1991) 423.

[95] T. Falk, R. Madden, K.A. Olive, and M. Srednicki, *Phys. Lett.* **B318** (1993) 354.

[96] T. Falk, K.A. Olive, and M. Srednicki, *Phys. Lett.* **B354** (1995) 99; T. Falk and K.A. Olive, *Phys. Lett.* **B** in press, hep-ph/9602299 (1996).

[97] J. McDonald, K. A. Olive and M. Srednicki, *Phys. Lett.* **B283** (1992) 80.

[98] K. Greist, M. Kamionkowski, and M.S. Turner, *Phys. Rev.* **D41** (1990) 3565.

[99] M. Drees and M.M. Nojiri, *Phys. Rev.* **D47** (1993) 376.

[100] J. Lopez, D.V. Nanopoulos, and K. Yuan, *Nucl. Phys.* **B370** (1992) 445.

[101] K. Greist and D. Seckel, *Phys. Rev.* **D43** (1991) 3191; S. Mizuta and M. Yamaguchi, *Phys. Lett.* **B298** (1993) 120.

[102] J. Lu, these proceedings.

[103] M. Khlopov, these proceedings.

[104] P. Gondolo, these proceedings.

[105] L. Bergstrom, these proceedings.

[106] B. Cabrera, these proceedings.

[107] L. Mosca, J. Ramachers, V. Chazal, L. Zerle, T. Shutt, M. Pavan, K.I. Fushini, T. Ali, D. Snowden-Ifft, B. Lanou, these proceedings.

[108] R. Barbieri, M. Frigeni, and G.F. Guidice, *Nucl. Phys.* **B313** (1989) 725.

[109] R. Flores, K.A. Olive and M. Srednicki, *Phys. Lett.* **B237** (1990) 72.

[110] J. Ellis and R. Flores, *Nucl. Phys.* **307** (1988) 883; *Phys. Lett.* **B263** (1991) 259.

[111] L. Bergstom and P. Gondolo, hep-ph/9510252 (1995).

[112] J. Silk and M. Srednicki, *Phys. Rev. Lett.* **53** (1984) 624.

[113] J. Silk, K.A. Olive, and M. Srednicki, *Phys. Rev. Lett.* **55** (1985) 257.

[114] L. Krauss, M. Srednicki, and F. Wilczek, *Phys. Rev.* **D33** (1986) 2079; K. Freese, *Phys. Lett.* **B167** (1986) 295.

[115] J. Edsjo, these proceedings.

[116] A. DeRujula, these proceedings.

[117] G. Tarle, these proceedings.

WEAK LENSING AS A PROBE OF DARK MATTER

L. Van Waerbeke[1], Y. Mellier[2]
[1] *Observatoire Midi Pyrénées,*
14 Av. Edouard Belin,
31400 Toulouse, France
[2] *Institut d'Astrophysique de Paris*
98bis Boulevard Arago
75014 Paris, France

Abstract

Weak gravitational lensing of distant galaxies can probe the total projected mass distribution of foreground gravitational structures on all scales and has been used successfully to map the projected mass distribution of rich intermediate redshift clusters. This paper reviews the general concepts of the lensing analysis. We focus on the relation between the observable (shapes and fluxes) and physical (mass, redshift) quantities and discuss some observational issues and recent developments on data analysis which appear promising for a better measurement of the lensing signatures (distortion and magnification) at very large scales.

1 Introduction

The dark matter (DM) component of gravitational structures is extensively studied from the dynamical analysis of the luminous component. Popular examples are the rotation curves of galaxies, motions of galaxies in groups or in clusters, or large scale velocity fields from which the mass and the distribution of the DM can be inferred, provided one assumes a dynamical state (Virial state) and a geometry (sphericity) of the gravitational system. Unfortunately, in most cases these hypothese are not fulfilled: for example, the Virial hypothesis applied to clusters may be wrong, because clusters may be young gravitational objects. Their mass profile could be alternatively obtained from the X-ray Bremsstralung emission of their intra-cluster gas which depends on their total mass distribution and their equilibrium state as well. Again, one has to assume a geometry and a thermodynamical state for the gas of photons and electrons. Despite these difficulties, all these studies provide similar trends, with the mass to light ratio M/L increasing with scale. For a typical galaxy M/L ranges between $10-30$, but is roughly 10 times larger for a cluster from both Virial and X-ray studies, leading to $\Omega \sim 0.02$ for the former and $\Omega \sim 0.2$ for the later. This means that the mass cannot be concentrated only within the central visible parts of galaxies.

Gravitational lensing provides a direct measurement of the projected mass density without additional hypothesis on the dynamical state or on the geometry of the mass distribution. Provided that we can measure the optical distortion of background objects caused by a foreground

mass, it is possible to constrain the projected mass distribution of this deflector. Recent results of the lensing analysis on some clusters are summerized in Table (1).

MASS DISTRIBUTION IN CLUSTERS OF GALAXIES

FROM WEAK LENSING

CLUSTER	Z	SCALE	M/L	Z(SOURCE)	REF.
1455+22	0.26	500 Kpc	460		Smail et al. 1994
Cl0016+16	0.55	500 Kpc	430		Smail et al. 1994
MS1224	0.33	500 Kpc	800	1.0 -- 2.0	Fahlmann et al. 1994
Cl0024+17	0.39	2.5 Mpc	600	0.9 -- 1.2	Bonnet et al. 1994
A1689	0.18	1.0 Mpc	400	1.0 -- 2.0	Tyson & Fischer 1995
A2218	0.18	400 Kpc	440	1.0 -- 2.0	Squires et al. 1995
A2390	0.23	1 Mpc	320	1.0 -- 2.0	Squires et al. 1996
MS1054	0.89	1.9 Mpc	1600 IF z<1 580 IF z=1.5 350 IF z=3		Luppino & Kaiser 1996
A1689	0.18	500 Kpc	>200	1.0 -- 2.0	Broadhurst 1996
A1689	0.18	1 Mpc	400	1.0 -- 2.0	Kaiser 1996
Cl0939	0.41	400 Kpc	200	0.6 -- 1.0	Seitz et al. 1996

Table 1: Status of the observations.

These results are larger than the usual Virial or X-Ray analysis by a factor 2 or 3. Whether this discrepancy may be explained or not is not clear yet (Miralda-Escudé & Babul (1995), Navarro et al. (1995)), and we do not discuss this here. We discuss in this paper the general method of the weak lensing analysis leading to these results.

2 The weak lensing analysis

2.1 Basics of gravitational lensing

Because of gravitational lensing, ray-lights are bended and the apparent position $\vec{\theta}_I$ differs from the source position in absence of lensing $\vec{\theta}_S$ by the quantity $\vec{\alpha}$:

$$\vec{\theta}_I = \vec{\theta}_S + \vec{\alpha} \tag{1}$$

where $\vec{\alpha}$ is the gradient of the two-dimensional (projected on the line of sight) gravitational potential ϕ. The gravitational distortion of background objects is described by the Jacobian of the transformation, namely the amplification matrix $\mathcal{A}$ between the source and the image plane (for more details, see Schneider, Ehlers & Falco 1992):

$$\mathcal{A} = \begin{pmatrix} 1-\kappa-\gamma_1 & -\gamma_2 \\ -\gamma_2 & 1-\kappa+\gamma_1 \end{pmatrix} \tag{2}$$

where κ is the convergence, γ_1 and γ_2 are the shear components, and are related to the Newtonian gravitational potential ϕ by:

$$\kappa = \frac{1}{2}\nabla^2\phi = \frac{\Sigma}{\Sigma_{crit}} \tag{3}$$

$$\gamma_1 = \frac{1}{2}(\phi_{,11} - \phi_{,22}) \; ; \quad \gamma_2 = \frac{1}{2}\phi_{,12} \; , \tag{4}$$

where Σ is the projected mass density and Σ_{crit} is the critical mass density which depends on the angular diameter distances D_{ij}, $(i,j = (o(bserver), l(ens), s(ource)))$ involved in the lens configuration:

$$\Sigma_{crit} = \frac{c^2}{4\pi G}\frac{D_{os}}{D_{ol}D_{ls}} \tag{5}$$

The quantities κ, γ_1, and γ_2 are not observables. Only the magnification μ and the distortion δ are in principle observable quantities because they are related to the fluxes and the shapes of the objects (Miralda-Escudé 1991, Schneider & Seitz 1995):

$$\mu = \frac{1}{|\mathcal{A}|} = \frac{1}{(1-\kappa)^2 - \gamma^2} \tag{6}$$

$$\delta_i = \frac{2g_i}{1+|g|^2} \; ; \; g_i = \frac{\gamma_i}{1-\kappa} \tag{7}$$

Since we are interested in the large scale distribution of the Dark Matter ($> 0.5 Mpc$) we only focus the analysis on the weak lensing regime where $(\kappa, \gamma) << 1$. The relations between the physical (γ, κ) and observable (δ, μ) quantities become more simple:

$$\mu = 1 + 2\kappa \tag{8}$$

$$\delta_i = 2\gamma_i \tag{9}$$

The projected mass density Σ of the lens is available from the amplification μ using Eq. (8) and Eq. (3), or equivalently from the distortion field by using Eq.(9) and the integration of Eq. (4) (Kaiser, Squires (1993)):

$$\kappa(\vec{\theta}_I) = \frac{-2}{\pi}\int d\vec{\theta}\frac{\chi_i(\vec{\theta}-\vec{\theta}_I)\gamma_i(\vec{\theta}_I)}{(\vec{\theta}-\vec{\theta}_I)^2} + \kappa_0 \; ; \quad \vec{\chi}(\vec{\theta}) = (\frac{\theta_1^2-\theta_2^2}{\theta^2} \; , \; \frac{2\theta_1\theta_2}{\theta^2}) \tag{10}$$

κ_0 is the integration constant. In the weak lensing regime, Eqs.(8) and (9) provide two independant methods to map the total projected mass, using the distortion of the background objects and the magnification of the background objects.

2.2 How observable quantities are measured?

The gravitational distortion is not visible on a single galaxy in the weak lensing regime because $\gamma << \bar{\epsilon}$, where $\bar{\epsilon}$ is the mean ellipticity of the galaxies. Fortunately a gravitational shear in a given area of the sky distorts all the background galaxies by a same amount, and the distortion can be measured from the mean polarization of these galaxies. The distortion δ_i is computed from the shape of the galaxies in the image plane. Each galaxy is assumed to be elliptical with an ellipticity ϵ and an orientation θ, and is described by a polarization vector $\epsilon e^{i\theta}$ in a complex formalism. The distortion is given by the sum of the polarization vectors in a given area of the sky. No information is required in the source plane, only the isotropy of the orientation of the galaxies in the source plane is assumed. A detailled description on the optimum detection and measurement of the shape of the galaxies is given in Bonnet & Mellier (1995) and Kaiser et al. (1995). This method has been succesfully applied to several clusters (See Table(1)), and to simulations to get the distortion. However, the intrinsic ellipticity of the galaxies is a source of noise, and the contribution of the random orientations of N_g galaxies to the value of the shear γ_i is given by $\bar{\epsilon}/\sqrt{N_g}$. An estimate of the mass using Eq. (10) requires an estimation of Σ_{crit}, which implies the redshift of the sources which is poorly known. Though this is not a critical issue for nearby clusters ($z_l < 0.2$) because $D_{os}/D_{ls} \simeq 1$, it could lead to a large uncertainty of the mass for more distant clusters (See Table (1)).

Unfortunately it is impossible to get a true value for the mass only from the shear map, even if we know the redshift of the sources, because a constant mass plane does not induce any shear on background galaxies. Mathematically, this corresponds to the unknown intregration constant κ_0 in Eq.(10). This degeneracy may be broken if one measures the magnification μ which depends on the mass quantity inside the light beam (Eq.(3)). While the shear measurement does not require any information in the source plane, the magnification measurement needs the observation of a reference (unlensed) field to calibrate the magnification. Broadhurst et al. 1995 proposed to compare the number count $N(m, z)$ and/or $N(m)$ in a lensed and an unlensed field to measure μ. Depending on the value of the slope S of the number count in the reference field, we observe a bias (more objects) or an anti-bias (less objects) in the lensed field. The particular value $S = 0.4$ corresponds to the case where the magnification of faint objects is exactely compensated by the dilution of the number count. This method was applied successfully on the cluster A1689 (Broadhurst, 1995), but the signal to noise of the detection remains 5 times lower than with the distortion method for a given number of galaxies. The magnification may also be determined by the changes of the image sizes at fixed surface brightness (Bartelmann & Narayan 1995).

The weakness of these methods is that they require to measure the shape, size and magnitude of very faint objects up to B=28, and this is not sure whether the measurement is optimum, and whether systematic effects are avoided. The determination of the shape parameter depends on the threshold level and the convolution mask, and in any cases the information contained in pixels fainter than the threshold level is lost. Furthermore, the measurement of the shape from the second moment matrix is equivalent to the assumption that the objects are elliptical, which is not true. These remarks lead us to propose a new and independent method to analyse the lensing effects, based on the auto-correlation function of the pixels in CCD images, which avoids shape parameter measurements (Van Waerbeke et al. 1996).

3 The Auto-correlation method

3.1 Principle

The CCD image is viewed as a density field rather than an image containing delimited objects. The surface brightness in the image plane in the direction $\vec{\theta}$ is related to the surface brightness in the source plane $I^{(s)}$ by the relation:

$$I(\vec{\theta}) = I^{(s)}(\mathcal{A}\vec{\theta}) \tag{11}$$

and for the auto-correlation function (ACF):

$$\xi(\vec{\theta}) = \xi^{(s)}(\mathcal{A}\vec{\theta}) \tag{12}$$

To understand the meaning of this equation, let us write it in the weak lensing regime:

$$\xi(\vec{\theta}) = \xi^{(s)}(\theta) - \theta\ \partial_\theta \xi^{(s)}(\theta)[1 - \mathcal{A}] \tag{13}$$

$\xi(\vec{\theta})$ is the sum of an isotropic unlensed term $\xi^{(s)}(\theta)$, an isotropic lens term which depends on κ, and an anisotropic term which depends on γ_i.

Let us analyse which gravitational lensing information can be extracted from the shape matrix $\mathcal{M}$ of ξ:

$$\mathcal{M}_{ij} = \frac{\int d^2\theta \xi(\vec{\theta})\theta_i\theta_j}{\int d^2\theta \xi(\vec{\theta})} \tag{14}$$

The shape matrix in the image plane is simply related to the shape matrix in the source plane $\mathcal{M}^{(s)}$ by $\mathcal{M}_{ij} = \mathcal{A}^{-1}_{ik}\mathcal{A}^{-1}_{jl}\mathcal{M}^{(s)}_{kl}$. If the galaxies are isotropically distributed in the source plane, $\xi^{(s)}$ is isotropic, and in that case $\mathcal{M}^{(s)}_{ij} = M\delta_{ij}$, where δ_{ij} is the identity matrix. Using the expression of the amplification matrix $\mathcal{A}$ we get the general form for $\mathcal{M}$:

$$\mathcal{M} = \frac{M(a + |g|^2)}{(1-\kappa)^2(1-|g|^2)} \begin{pmatrix} 1+\delta_1 & \delta_2 \\ \delta_2 & 1-\delta_1 \end{pmatrix} \tag{15}$$

The observable quantities (distortion δ_i and magnification μ) are given in terms of the components of the shape matrix:

$$\delta_1 = \frac{\mathcal{M}_{11} - \mathcal{M}_{22}}{tr\mathcal{M}}\ ; \quad \delta_2 = \frac{2\mathcal{M}_{12}}{tr\mathcal{M}}\ ; \quad \mu = \sqrt{\frac{det\mathcal{M}}{M}} \tag{16}$$

where $tr\mathcal{M}$ is the trace of $\mathcal{M}$. As for lensed galaxies, we see that the distortion is available from a direct measurement in the image plane while the magnification measurement requires to know the value of M which is related to the light distribution in the source plane, or in an unlensed reference plane. The ACF provides a new and independent way to measure δ_i and μ which does not require shape, size or photometry of individual galaxies. In the following we only describe the measurement of the distortion using this method. The case of the magnification which requires an analysis of the sources galaxies in a reference field, will be developped in a future work.

3.2 The practical method

By definition, the value of the ACF at a pixel position ij is $E_{ij} = \frac{1}{N_{pix}} \sum_{kl} (I_{i+kj+l} - \bar{I})(I_{kl} - \bar{I})$, where I_{kl} is the value of the pixel kl, $\bar{I}$ is the mean value of the image, and N_{pix} the number of terms in the sum. The ACF is computed in a part of the image (a superpixel) where the shear is assumed to be constant in intensity and direction. Two strategies are possible to compute the ACF. First we can remove all the unwanted objects (stars, bright galaxies, dead CCD lines, cosmetic defaults,...) and compute the ACF from the rest of the image. The main interest of this approach is that it works at the noise level and even ultra faint objects are taken into account. The second approach consists in selecting objects from a given criteria (magnitude, colors, redshift,...), in surrounding them by a large circle, put the rest of the image to zero and compute the ACF of the image containing these circles.

As for the case of individual galaxy we need to compute the shape matrix of the ACF in an annular filter (Bonnet & Mellier, 1995) to avoid the center, where the signal is strongly polluted by the Point Spread Function (PSF), and the external part, which is dominated by the noise. The effects of the PSF and the filter are calibrated by using simulations.

3.3 Sources of errors

The galaxies have not the same flux, size and profile and, by definition of the ACF, are weighted by the square of their flux. Since this could change the statistical properties of the ACF, it is better to work with selected objects by using the second strategy of the ACF method. The idea is to weight each circle which contains an object by a multiplicative term defined as $[\frac{1}{N_{pix}} \sum_{ij} (I_{ij} - \bar{I})^2]^{-1/2}$, where N_{pix} is the number of pixels of the object. The objects are then equally weighted, even when they have very different magnitudes, sizes and profiles.

The intrinsic ellipticity of the galaxies induces a statistical dispersion on the shear estimate of $\sqrt{\bar{\epsilon}/N_g}$, where $\bar{\epsilon}$ is the mean ellipticity of N_g galaxies. Instrumental errors, tracking errors or anisotropic PSF may be removed provided they are measurable on the stellar profiles (Bonnet & Mellier 1995, Kaiser et al. 1995).

The photon noise is a source of error of this method. Indeed, the distortion is computed from one object, the ACF itself. Since the noise polarizes randomly an object, a high noise level makes the measurement of the weak distortion impossible. This lead to the conclusion that a given level of noise corresponds to a distortion threshold γ_0 below which the measured distortion is not reliable. We quantified this threshold from simulations.

4 Conclusion

An optimum analysis of the lensing effects requires the measurement of both the distortion and the magnification to comfirm and improve the results quoted in Table (1), and to measure the very weak shear caused by large scale structures. Van Waerkebe et al. (1996) have proposed a new and independent method to measure the gravitational distortion of the background galaxies from the auto-correlation function of the brightness distribution. It does not require any shape, size and centroid determination of individual galaxies, and avoids possible systematics. Moreover the resulting shear is unique and does not depends on the choice of the detection criteria.

The method has been checked on simulated and true data (Q2345 and CL0024). An example of the shear analysis using the ACF on simulated data is shown on Figures 1,2. The shear maps

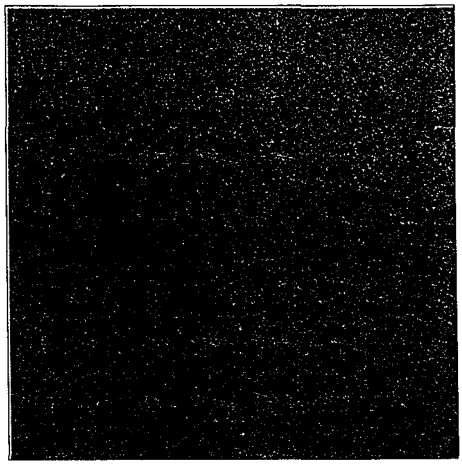


Figure 1: Simulation of a 4 hours exposure at CFHT in the B band on a 3.5" × 3.5" field. The seeing is 0.7" with no tracking errors. Galaxies are lensed by an isothermal sphere (σ = 1000km/s), with a core radius of 4″ located at 200″ bottom from the field center. The lens redshift is 0.17 and the mean redshift of the sources is 1. The segments show the local orientation of the shear. Their length is proportional to the shear intensity.

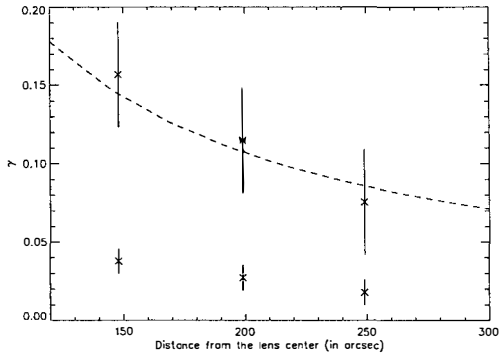


Figure 2: 1-dimensional shear profile from the simulation of Fig.1. At the bottom the uncalibrated measure points are drawn. The theoretically expected shear profile is plotted as the dashed line.

of the real images Q2345 and CL0024 were previously obtained by Bonnet et al. (1993,1994) with the standard method of individual galaxy analysis. Our results are in very good agreement. Moreover, because of the increase of the sensibility with our method, we predict the existence of a new gravitational deflector in the field of Q2345. Further observations will check this point. Because of its simplicity and robustness, this method is well adapted to measure weak shear caused by large scale structure for which a large number of galaxies ($\sim$ 100000) is required.

Acknowledgements. We thanks P. Schneider, F. Bernardeau and B. Fort for discussions and enthusiastic support. LVW thanks B. Guiderdoni for his invitation and the Moriond's staff for hospitality and financial support.

5 References

Bartelmann, M., Narayan, R. (1995) ApJ 451, 60.
Bonnet, H., Fort, B., Kneib, J-P., Mellier, Y., Soucail, G. (1993) A&A 280, L7
Bonnet, H., Mellier, Y., Fort, B. (1994) ApJ 427, L83.
Bonnet, H., Mellier, Y. (1995) A&A 303, 331.
Broadhurst, T. (1996) SISSA preprint astro-ph/9511150.
Broadhurst, T., Taylor, A.N., Peacock (1995) ApJ 438, 49.
Falhman, G., Kaiser, N., Squires, G., Woods, D. (1994) ApJ 437. 56.
Kaiser, N. (1996) SISSA preprint astro-ph/9509019.
Kaiser, N., Squires, G. (1993) ApJ 404, 441
Kaiser, N., Squires, G., Broadhurst, T. (1995) ApJ 449, 460.
Luppino, G., Kaiser, N. (1996) SISSA preprint astro-ph/9601194.
Miralda-Escudé, J., Babul, A. (1995) ApJ, 449, 18.
Navarro, J., Frenk, C., White, S. (1995) MNRAS, 275, 720.
Schneider, P., Ehlers, J., Falco, E. E., (1992), *Gravitational Lenses*, Springer.
Schneider, P., Seitz, C. (1995) A&A 294, 411.
Seitz, C., Kneib, J.P., Schneider, P., Seitz, S., (1996) in press.
Smail, I., Ellis, R.S., Fitchett, M. (1994) MNRAS 270, 245.
Squires, G., Kaiser, N., Babul, A., Fahlmann, G., Woods, D., Neumann, D.M., Böhringer, H. (1995) submitted.
Squires, G., Kaiser, N., Falhman, G., Babul, A., Woods, D. (1996) SISSA preprint astro-ph/9602105.
Tyson, J.A., Fisher, P. (1995) ApJL, 349, L1.
Van Waerbeke, L., Mellier, Y., Schneider, P., Fort, B., Mathez, G. (1996) A&A in press.

PARTICLE DARK MATTER REVIEW

Gerard Jungman
Department of Physics, Syracuse University
Syracuse, NY 13244 USA

Abstract

I give an introduction to particle dark matter models. The emphasis is on particle physics models that give cold dark matter. The case against light neutrinos is recalled. Axions and neutralinos from supersymmetry are discussed.

1 Baryons

Before considering nonbaryonic dark matter, something should be said about baryons in the Universe. The baryonic content of the Universe and baryonic dark matter are discussed in greater detail in the contribution of K. Olive elsewhere in these proceedings.

Our knowledge of the total baryonic content of the Universe comes from calculations of big-bang nucleosynthesis. It turns out that, in order to reproduce the observed primordial elemental abundances, the primordial baryon to photon ratio η must be fixed in the range $3 \times 10^{-10} \lesssim \eta \lesssim 8 \times 10^{-10}$, and this can be translated into a range for the baryonic content in units of the critical density $0.01 \lesssim \Omega_b h^2 \lesssim 0.03$. The upper bound limits the baryonic contribution to the present density of the Universe, and provides an important argument for nonbaryonic dark matter.[1] Of course the details of the story are not quite so straightforward. Much work goes into understanding systematic effects in the interpretation of the elemental abundance data. I refer again to the contribution of K. Olive.

It is worth pointing out that a precision map of microwave background fluctuations should provide an independent constraint on Ω_b in the future.[2]

2 Light Neutrinos

In the standard model of particle physics, the three neutrino species are taken to be exactly massless. The direct experimental limit[3] on the electron neutrino mass is approximately 5 eV, but the limits on the muon and tau neutrinos are much less stringent; $m_{\nu_\mu} \lesssim 270$ keV and

$m_{\nu_\tau} \lesssim 31$ MeV. Although these masses are quite small compared to the Dirac masses of the associated charged leptons, this smallness is a natural state of affairs in many extensions of the standard model. Many grand unified models require ultra-massive right-handed neutrino states and lepton-number violation at high energy scales. These effects can combine to give the observed neutrinos small masses through the seesaw mechanism.[4,5]

The cosmological consequences of neutrino masses of order 0.1 keV have been understood for some time. Weak interactions keep the neutrinos in thermal equilibrium above temperatures of order 1 MeV. Below this temperature the neutrino abundance freezes out, but the distribution remains thermal as the particle energy and momenta redshift in the expanding universe. These relic neutrinos will contribute an energy density to the universe at the present epoch which can be non-negligible if they have sufficient mass. One finds a relic abundance in units of the critical density of order[6,7] $\Omega_\nu h^2 \simeq m_\nu/90$ eV. Therefore, massive but light neutrinos can be an interesting dark matter candidate. If one allows the neutrinos to decay, a more interesting situation can develop.[8]

However, there is more to the story than simply counting the relic neutrinos. Several convincing arguments show that neutrinos cannot solve the observed dark matter problem. The first of these depends on our understanding of large-scale structure formation. Lighter particles become nonrelativistic at later times and lower temperatures, and a relativistic collisionless gas will erase density perturbations on small enough scales, by free streaming. Thermal relic light neutrinos will still be relativistic at the time when objects of galactic size, approximately $10^{12}\ M_\odot$, should be starting to collapse. Therefore, the neutrinos will begin to take part in gravitational collapse only when the horizon is much larger, encompassing objects of the size of superclusters, approximately $10^{15}\ M_\odot$. Such behaviour classifies these light neutrinos as "hot dark matter". It turns out to be quite difficult to build models dominated by hot dark matter which produce the observed distribution of matter in the Universe. It appears instead that a hot dark matter component can be at most a perturbation on a "cold dark matter" background model, where massive nonrelativistic particles can collapse directly to form objects of galactic size.[9] Of course, it should be kept in mind that structure formation based on topological defect seeds may change these conclusions.[10]

Furthermore, there is direct observational evidence for a component of dark matter in galaxies which cannot be light neutrinos. Consider the population of light neutrinos in a galactic halo, which formed when neutrinos relaxed into their own gravitational well. Because the neutrino gas is collisionless, the phase space density is conserved along particle trajectories. Therefore, if we write the halo velocity distribution in the form

$$f_\nu(v) = n_0 \exp\left[-\frac{1}{2}v^2 + \phi\right], \tag{1}$$

where ϕ is the gravitational potential, then the constant n_0 can be directly related to the distribution at early times when the gravitational potential was negligible. Since that early-time distribution was of a thermal form, it must satisfy an inequality stemming from $(1 + \exp(E/T))^{-1} \leq 0.5$. Of course this is just the manifestation of the Pauli principle. Therefore the total occupation of phase space in the halo, which is obtained by integrating $f_\nu(v)$ over velocities, has an inherent upper bound which depends on the form of the halo.[11] Now, there exist examples of very compact galaxies with large dark halos, such as the dwarf spheroidal galaxies Draco and Ursa Minor. If the halos of these objects are composed entirely of light neutrinos then these neutrinos must have large masses in order to avoid the phase space bound; $m_\nu > 400$ eV.[12,13] But such masses are not acceptable because they would violate the upper bound coming from the relic density calculation. Therefore we conclude that the halos of these

objects must have some large component which is not light neutrinos.

These and other considerations have lead us away from hot dark matter models and toward cold dark matter models. It turns out that many extensions of the standard model give cold dark matter candidates. Several of these have become canonical examples, and we discuss them now.

3 WIMPs

3.1 Generalities

As mentioned above, a Universe dominated by light thermal relics will have difficulty producing the observed distribution of galaxies. Therefore it is a natural step to consider hypothetical weakly interacting particles which are heavy enough to be nonrelativistic at the start of the epoch of galaxy formation. Such particles are called WIMPs (weakly-interacting massive particles).

Consider such a hypothetical stable massive particle. Suppose that pairs of these particles can annihilate with each other and that this annihilation process insures that the particles were in thermal equilibrium at early times, which will be true in the case at hand. The kinetic equation for the evolution of the density of these particles in the expanding Universe is

$$\dot{n} + 3\frac{\dot{a}}{a}n = \langle\sigma v\rangle(n_{\rm eq}^2 - n^2), \tag{2}$$

where $n_{\rm eq}$ is the equilibrium density at the given temperature and $\langle\sigma v\rangle$ is the average over the velocity distribution of the annihilation cross section times velocity of the particles.

At early times the rate of the annihilation process will be much faster then the rate of expansion of the Universe, and thermal equilibrium will result. At some later time, when the density has been diluted by the expansion, the expansion rate can begin to dominate the annihilation rate and the particle density will effectly freeze-out at some (comoving) density value. The resulting relic density at the present epoch can be computed directly from the above evolution equation.[14] For an estimate we can make the rough approximation that $\langle\sigma v\rangle$ is constant. In that case the relic density will be $\Omega_W h^2 \simeq 3 \times 10^{-10}\,{\rm GeV}^{-2}/\langle\sigma v\rangle$. Notice the interesting numerical coincidence; the dimensionful parameter appearing here is a *weak-scale* cross section. Therefore, if the particle has an annihilation cross section of weak strength then it will naturally have a relic abundance of order unity. This is what makes WIMP dark matter natural from the stand-point of many particle physics models which are relevant to electroweak symmetry breaking and the electro-weak interactions.

3.2 Heavy Neutrinos

Perhaps the simplest example of a WIMP is a heavy neutrino; for definiteness one can think of a Dirac neutrino with a mass of order tens of GeV.[15–17] We will assume that this neutrino has the same electroweak charges as the known neutrinos. Such an object could arise from a hypothetical fourth generation. Since we fix all the couplings by assumption, there is really only one free parameter, which is the mass. This makes the model very simple to analyze. The relic abundance[15–17] turns out to be $\Omega_\nu h^2 \simeq (m_\nu/5\ {\rm GeV})^{-2}$.

All appears well for the heavy neutrino candidate. However, it is perhaps the first major success story of dark matter direct detection efforts that such a particle is essentially ruled out as the major component of our galactic halo. Because the couplings and interactions of such

a particle are fixed by assumption, it is possible to calculate the rate of interaction of such halo WIMPs in low-background cryogenic detectors. The result is that such particles, if they constituted the major fraction of the Milky Way halo, would have to weigh in excess of roughly 4 TeV.[18,19] But if they weighed this much, then their relic abundance would be too small for them to consistently dominate galactic halos. It is important here that a lower mass window is closed by direct search at LEP. For more details of direct dark matter detection efforts, see the contribution of P. Gondolo to these proceedings.

3.3 Neutralinos

As mentioned above, theories relevant to electroweak symmetry breaking which introduce extra stable particles will generally give WIMP dark matter candidates. Supersymmetry is precisely such a theory. One of the main motivations for supersymmetry is to understand how a scale as small as the weak scale could make dynamical sense in a unified theory where the "natural" scale for physics might be $M_{GUT} \simeq 10^{16}$ GeV, or even $M_{Pl} \simeq 10^{19}$ GeV if one considers possible unification with gravity. The small ratio M_{weak}/M_{GUT} needs either to be "fine-tuned" by hand or stabilized against upsetting dynamical effects. Supersymmetry introduces fermionic partners for the known bosons, and vice versa, and the effects of these new particles are such as to exactly cancel the badly behaved radiative corrections which have the potential to erase the required small ratio of scales. This cancellation is exact in the limit that supersymmetry is exact.

Since the world is not manifestly supersymmetric, some supersymmetry breaking must be introduced. Although there are many possible mechanisms for supersymmetry breaking, and the subject is very much an open field of interest, one thing we can say for sure is that the scale of the supersymmetry breaking terms should not be much larger than the weak-scale, otherwise we would loose the explanation for the stability of the small ratio M_{weak}/M_{GUT}. Therefore, in particular, the masses of the new superpartner particles should not be much larger than the weak scale, and some of these particles will be weakly interacting since they are partners of normal weakly interacting particles. Therefore we expect to find WIMP dark matter candidates in supersymmetric models.

Consider the minimal application of supersymmetry to the standard model.[20,21] The symmetry will give rise to partners of quarks, leptons, Higgs bosons, and gauge bosons. Note that for technical reasons there must be at least two Higgs bosons in the model. A kind of discrete relic of supersymmetry called R-parity will insure that the lightest superpartner particle will be absolutely stable; we will assume that R-parity is unbroken so that this is true. Otherwise our candidate relic superpartner particle would likely decay and thus not be of interest. The lightest superpartner particle, denoted LSP, is then our WIMP candidate.

Sleptons, the superpartners of leptons, are not good dark matter candidates because a relic abundance of massive charged particles would be very difficult to reconcile with observations. At the very least we would expect to see such particles in the primary cosmic rays, and none are seen.[22,23]

Stable squarks would presumably bind into color-neutral hadronic objects. They could be arranged to be electrically neutral, and could make an interesting candidate. However, they would be strongly interacting and so the physics would be different from the canonical WIMP models discussed here. Furthermore, there is a theoretical prejudice against this occurence, coming from grand-unified models. Assuming grand-unification, it turns out that strong-interaction radiative corrections will generically drive the squark masses higher than the slepton masses and so a squark could not be the LSP.

Sneutrinos would be quite interesting candidates.[24] However, the limits on direct detection of heavy neutrinos apply similarly to sneutrinos because the cross-sections for sneutrino and neutrino elastic scattering from nuclei are directly related. The result is that sneutrino masses below $\mathcal{O}(1\text{ TeV})$ are ruled out.[25]

This leaves the partners of the gauge and Higgs bosons, $\tilde{Z}$, $\tilde{\gamma}$, $\tilde{H}_1$, $\tilde{H}_2$. The partners of these particles will be heavy, electrically neutral, and weakly interacting. The actual physical states will be mixtures of the above due to the combined effects of electroweak symmetry breaking and supersymmetry breaking. We call these physical states the *neutralinos*, denoting them by χ_i, $i = 1, 2, 3, 4$. The lightest of these will be our WIMP dark matter candidate.

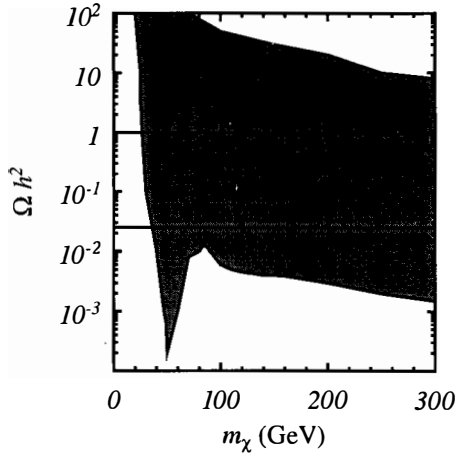


Figure 1: Neutralino relic abundance for a range of viable models of minimal supersymmetry. See Ref. 21 for details.

A detailed analysis of neutralinos in supersymmetric models requires some effort because the models can be tediously complicated. As an example, Fig. 1 shows the results for neutralino relic abundance in a large class of viable models. The horizontal bars indicate the cosmologically interesting range for the relic abundance. I refer to Ref. 21.

4 Axions

Up until now we have discussed thermal relics, particles which were in thermal equlibrium at some early time and which therefore have a thermal form of distribution at later times. This allowed many calculations to be carried out without detailed dynamical information. However, we can also consider models where the candidate dark matter particle was produced in some way other than through thermal freeze-out. Axions are a canonical such candidate.

Axions are motivated in particle physics by the desire to obtain a dynamical explanation of the so-called strong-CP- or Θ-problem. This problem can be explained in the following way. It is possible to introduce into the QCD Lagrangian the following interaction term

$$\mathcal{L}_\theta = \frac{\theta}{8\pi} \int tr(G_{\mu\nu}\tilde{G}^{\mu\nu}). \tag{3}$$

This term is of dimension four and should generically not be suppressed by any small mass ratios or other parameters. Classically it is irrelevant since it is a total derivative which does not effect the equations of motion. However, this term will effect the quantum theory, and it has the curious property that it breaks P and CP symmetries.

One can estimate what effect this term would have on low-energy hadron physics. The most disturbing prediction is the existence of an electric dipole moment for the neutron, which leads to the limit $\theta \lesssim 10^{-9}$. One is left to explain why this coupling constant should be so small. The idea of Peccei and Quinn was to elevate this constant to the status of a dynamical

field. They introduced a new $U(1)_{PQ}$ symmetry which could be spontaneously broken, leaving a Nambu-Goldstone boson, a, which would have a coupling to gluons in precisely the form $a\,tr(G_{\mu\nu}\tilde{G}^{\mu\nu})$, and θ would thus become dynamical. The effective θ coupling would then relax to zero when the new field variable a adjusts itself to the minimum of its effective potential.

The interactions of the axion field with ordinary matter depend on how the symmetry $U(1)_{PQ}$ is realized on the quarks and leptons. However, once the actions of the symmetry are specified the interactions are all determined by standard techniques for the effective interactions of Nambu-Goldstone bosons in the low-energy limit (current-algebra). So the models fall into classes which are specified by the action of $U(1)_{PQ}$ on the quarks and leptons.

Axions are not precisely massless because the $U(1)_{PQ}$ is not exact; it is broken by the same nonperturbative effects which break the axial $U(1)$ symmetry of QCD. One can think of the coupling to gluons as inducing a mixing between the axion and the pion. Therefore axions have a small mass, $m_a \simeq m_\pi f_\pi / f_{PQ}$, where m_π and f_π are the pion mass and decay constant respectively, and f_{PQ} is the "axion decay constant" which is related to the scale of $U(1)_{PQ}$ breaking. The interactions of axions are suppressed by powers of f_{PQ} as well. It was realized early that f_{PQ} must be large in order that the axion not be ruled out by laboratory experiments.[26,27]

Axion models have the important feature that they are relatively constrained. The models are characterized essentially by two parameters, one parameter being the axion mass and another being a coupling or mixing angle which depends on the type of model.[28–30]

There are three main physical effects which must be considered in the axion picture. These are axion production by relaxation of an initial "inflationary bubble" value, axion production by emission from global axionic strings, and the the effect of axions on astrophysical cooling processes. We will discuss these in turn.

Even though axions have a small mass, they actually provide a subtle form of cold dark matter. As the temperature lowers below Λ_{QCD}, the effective potential for the axion field turns on and the field begins oscillating around its minimum, with an amplitude that depends on its initial value. These oscillations of the field are interpretted quantum mechanically as particles, but any three-momentum of these particles is rapidly redshifted away, leaving only energy density. This oscillating field then corresponds to a zero-momentum condensate of bosonic axion particles.[31,32] That is why relic axions are cold, even though their mass can be very small. Note that the relic density depends on an assumption about the initial value of the axion field, which is a random variable, presumed constant throughout the "inflationary bubble" that we inhabit. The predicted relic axion density will depend on the value of this constant, Θ_i. We have[14,28,29]

$$\Omega_a h^2 \simeq 0.8 \times 10^{\pm 0.4} \Theta_i^2 \left(\frac{m_a}{10^{-5}\ \mathrm{eV}}\right)^{-1.18}. \tag{4}$$

One expects "typically" that Θ_i is of order unity.

There is another very interesting way that axions can be produced. At the Peccei-Quinn phase transition, $U(1)_{PQ}$ global cosmic strings will form. If the initial domains are not inflated, then strings will form within the horizon and will begin oscillating and decaying by emission of axion radiation. The resulting energy density of axions can be quite large.[33–35] Controversy existed for some time over the form of the radiation spectrum, but it appears that the literature has stabilized. The limit on the mass of the axion is approximately $m_a > 10^{-4}$ eV.[34]

Finally one must consider the effects of axion emission from the cores of astrophysical objects. Such emission provides a cooling mechanism which can compete with neutrino emission. Absence of anomalous cooling of red giants implies $m_a \lesssim 1$ eV. The observed SN1987a burst

extends this to $m_a \lesssim 10^{-3}$ eV except possibly for a window near 1 eV for "hadronic" axions, which have the most restricted couplings to known particles.[28,29]

Roughly speaking then, a target window for axion searches exists in the mass range $10^{-4}\text{eV} \lesssim m_a \lesssim 10^{-3}\text{eV}$, and possibly in a window near 1 eV for hadronic axions. The lower limit depends on the assumption that the string production mechanism operates. In the case that the initial domains were inflated, the lower limit cannot be considered rigorous, however a "typical" expectation could be taken as $m_a > 10^{-5}$ eV.

References

[1] T. P. Walker *et al.*, Astrophys. J. **376**, 51 (1991).

[2] G. Jungman, M. Kamionkowski, A. Kosowsky, and D. Spergel, astro-ph/9512139 (1995), to appear in Phys. Rev. D.

[3] M. Aguilar-Benitez and others (Particle Data Group), Phys. Rev. D **50**, 1177 (1994).

[4] M. Gell-Mann, P. Ramond, and R. Slansky, in *Supergravity* (North-Holland, Amsterdam, 1979).

[5] T. Yanagida, in *Proc. Workshop on Unified Theories and Baryon Number in the Universe* (KEK, Japan, 1979).

[6] R. Cowsik and J. McClelland, Phys. Rev. Lett. **29**, 669 (1972).

[7] G. Gerstein and Y. B. Zeldovich, JETP Lett. **4**, 174 (1966).

[8] D. A. Dicus, E. W. Kolb, and V. L. Teplitz, Phys. Rev. Lett. **39**, 168 (1977).

[9] G. R. Blumenthal, S. M. Faber, J. R. Primack, and M. J. Rees, Nature **311**, 517 (1984).

[10] R. Brandenberger, astro-ph/9508159 (1995).

[11] S. Tremaine and J. E. Gunn, Phys. Rev. Lett. **42**, 407 (1979).

[12] D. N. C. Lin and S. M. Faber, Astrophys. J. **266**, L21 (1983).

[13] M. Aaronson, Astrophys. J. **266**, L11 (1983).

[14] E. Kolb and M. S. Turner, *The Early Universe* (Addison-Wesley, Reading, 1989).

[15] B. W. Lee and S. Weinberg, Phys. Rev. Lett. **39**, 165 (1977).

[16] P. Hut, Phys. Lett. **B69**, 85 (1977).

[17] M. I. Vyotskii, A. D. Dolgov, and Y. B. Zeldovich, JETP Lett. **26**, 188 (1970).

[18] S. P. Ahlen *et al.*, Phys. Lett. **B195**, 603 (1987).

[19] D. O. Caldwell *et al.*, Phys. Rev. Lett. **61**, 510 (1988).

[20] H. E. Haber and G. L. Kane, Phys. Rep. **117**, 75 (1985).

[21] G. Jungman, M. Kamionkowski, and K. Griest, Phys. Rep. **267**, 195 (1996).

[22] S. Wolfram, Phys. Lett. **B82**, 65 (1979).

[23] P. F. Smith and J. R. J. Bennett, Nucl. Phys. **B149**, 525 (1979).

[24] T. Falk, K. A. Olive, and M. Srednicki, Phys. Lett. **B339**, 248 (1994).

[25] M. Beck *et al.*, Phys. Lett. **B336**, 141 (1994).

[26] J. E. Kim, Phys. Rev. Lett. **43**, 103 (1979).

[27] M. Dine, W. Fischler, and M. Srednicki, Phys. Lett. **B104**, 199 (1981).

[28] M. S. Turner, Phys. Rep. **197**, 67 (1990).

[29] G. Raffelt, Phys. Rep. **198**, 1 (1990).

[30] J. E. Kim, Phys. Rep. **150**, 1 (1987).

[31] J. Preskill, M. B. Wise, and F. Wilczek, Phys. Lett. **B120**, 127 (1983).

[32] L. F. Abbott and P. Sikivie, Phys. Lett. **B120**, 133 (1983).

[33] A. Vilenkin and E. P. S. Shellard, *Cosmic Strings and Other Topological Defects* (Cambridge University Press, Cambridge, 1994).

[34] R. A. Battye and E. P. S. Shellard, astro-ph/9412056 (1994).

[35] D. Harari and P. Sikivie, Phys. Lett. **B195**, 361 (1987).

PHENOMENOLOGICAL INTRODUCTION TO DIRECT DARK MATTER DETECTION

P. Gondolo
University of Oxford, Department of Physics, Theoretical Physics, 1 Keble Road, Oxford, OX1 3NP, United Kingdom

Abstract

The dark matter of our galactic halo may be constituted by elementary particles that interact weakly with ordinary matter (WIMPs). In spite of the very low counting rates expected for these dark matter particles to scatter off nuclei in a laboratory detector, such direct WIMP searches are possible and are experimentally carried out at present. An introduction to the theoretical ingredients entering the counting rates predictions, together with a short discussion of the major theoretical uncertainties, is here presented.

This is a phenomenological introduction to the detection of dark matter through its scattering in a laboratory detector. For dark matter in the form of massive quasi-stellar objects,

like brown dwarfs, which are much bigger and much heavier than the Earth, this type of detection is quite impracticable if not undesirable. I therefore consider dark matter in the form of elementary particles.

Many particles, most of which hypothetical, are at present candidates for dark matter: neutrinos, neutralinos, axions, etc. The methods employed in hunting for these particles are very different. In this short note I focus on this meeting's category of particle dark matter, *viz.* weakly interacting massive particles or WIMPs.

WIMPs, in a broad sense, are particles with masses of the order of atomic masses or higher ($m \gtrsim 10\text{GeV}/c^2$) that interact with ordinary matter with cross sections typical of the weak interaction or smaller ($\sigma \lesssim 10^{-38}\text{cm}^2$ off a proton). The presently most popular WIMP is the yet-undetected neutralino, the lightest supersymmetric particle in supersymmetric models. Other famous WIMPs are Dirac and Majorana neutrinos, which however, thanks to the on-going dark matter searches complemented by accelerator results, we know not to be the dominant component of our galactic halo.

A general introduction to dark matter has been given by Olive at this meeting. Direct detection of WIMPs was first explored by Goodman and Witten.[1] General reviews are Primack et al[2] and Smith and Lewin.[3] Engel *et al.*[4] present the nuclear physics involved. At this meeting Cabrera discusses experimental aspects of direct dark matter detection, while I focus on the theoretical aspects.

It is worth recalling some properties of the dark halo of our galaxy. Even if recent observations might change the details of our picture, the 1981 model by Caldwell and Ostriker[5] is good for my purposes. The Sun lies at a distance of ≈ 8.5 kpc on the disk of our spiral galaxy, and moves around the center at a speed of ≈ 220 km/s. The luminous disk extends to ≈ 12 kpc, and is surrounded by a halo of ≈ 100 kpc where globular star clusters and rare subdwarf stars are found. Dynamical arguments suggest that the halo is filled with dark matter, whose local density in the vicinity of the Sun is estimated to be $\rho_{\text{DM}} = 0.2\text{–}0.4\text{GeV}/c^2/\text{cm}^3 = 0.7\text{–}1.4\ 10^{-24}\text{g/cm}$.[3] Equilibrium considerations also give the root mean square velocity of halo consituents to be 200–400 km/s, not much different from the escape speed from the galaxy (500–700 km/s). Very little is known on the mean rotation speed of the halo, and we will assume it does not rotate. All in all, there is an optimistic factor of 2 uncertainty on the density and velocity of halo dark matter.

Are there WIMPs in our galactic halo? The scientific way to answer this question is to detect them. Signals could come from WIMP annihilation (indirect detection) or WIMP scat-

tering (direct detection). In the former we search for rare annihilation products like neutrinos, antimatter or gamma-ray lines. This is reviewed by Bergström at this meeting. In the latter, the basic philosophy is to build a target, wait and count.

The WIMP scattering rate per target nucleus is the product of the WIMP flux ϕ_χ and of the WIMP-nucleus cross section $\sigma_{\chi i}$. For an order of magnitude estimate we take the effective WIMP-nucleon coupling constant to be Fermi's constant $G_F = 2.3016\ 10^{-19}\hbar^2c^2\,\mathrm{cm/GeV}$, which sets the scale of weak interactions. We distinguish two cases: (i) the WIMP couples to nucleon spin, $\sigma_{\chi i} \approx G_F^2\mu_i^2/\hbar^4 \lesssim 10^{-34}\mathrm{cm}^2$; and (ii) the WIMP couples to nucleon number, $\sigma_{\chi i} \approx G_F^2\mu_i^2A_i^2/\hbar^4 \lesssim 10^{-30}\mathrm{cm}^2$. Here $\mu_i = m_\chi m_i/(m_\chi+m_i)$ is the reduced WIMP-nucleus mass, and A_i is the atomic number of the target (≈ 80 in the numerical examples) . The WIMP flux is $\phi_\chi = v\rho_\chi/m_\chi \approx 10^7\mathrm{cm}^{-2}\mathrm{s}^{-1}/(m_\chi c^2/\mathrm{GeV})$, for a WIMP density $\rho_\chi \approx 10^{-24}$ g/cm^3 and a typical WIMP velocity $v \approx 300$ km/s. The resulting scattering rates, taking $m_\chi \approx 100$ GeV/c^2, are of the order of $\lesssim 1$/kg-day for spin-coupled WIMPs and of $\lesssim 10^4$/kg-day for WIMPs coupled to nucleon number. These rates are quite small compared with normal radioactivity background. Therefore the common denominator of direct experimental searches of WIMPs is a fight against background.

For this we get help from characteristic signatures that we do not expect for the background. For example, while the Earth revolves around the Sun, the mean speed of the WIMP "wind" varies periodically with an amplitude of 60 km/s. This leads to a $\approx 10\%$ seasonal modulation in the detection rate, with a maximum in June and a minimum in December.[6] As another example, the direction of the WIMP "wind" does not coincides with the Earth rotation axis, so the detection rate might present a diurnal modulation due to the diffusion of WIMPs while they cross the Earth[7] (this however occurs for quite high cross sections). A final example of background discrimination is that the WIMP signal is directional, simply because most WIMPs come from the direction of the solar motion.[8]

WIMP-nucleus scattering

Since the relative speed $v \approx 300\mathrm{km/s} \approx 10^{-3}c$, the process can be treated non-relativistically. The center of mass momentum is given in terms of the reduced WIMP-nucleus mass as $k = \mu_i v$ and is $\lesssim A_i\,\mathrm{MeV}/c$ since $\mu_i \leq m_i$. The corresponding de Broglie wavelength is $\gtrsim 200\mathrm{fm}/A_i$, and can be smaller than the size of heavy target nuclei, in which case nuclear form factors are important. In the laboratory frame, the nucleus recoils with momentum $q = 2k\sin(\theta_{\mathrm{cm}}/2)$ and energy $\nu = q^2/2m_i$. Here θ_{cm} is the center-of-mass scattering angle. The 4-momentum transfer

is very small, $Q^2 \lesssim A_i^2\, 10^{-6}\,\mathrm{GeV}^2/c^2$ (compare with a typical deep inelastic $Q^2 \gtrsim 1\,\mathrm{GeV}^2/c^2$).

The differential scattering rate per unit recoil energy and unit target mass is formally

$$\frac{\mathrm{d}R}{\mathrm{d}\nu} = \frac{\rho_\chi}{m_\chi} \sum_i f_i \eta_i(q) \frac{\overline{|T_i(q^2)|^2}}{2\pi\hbar^4}. \tag{1}$$

The sum is over the nuclear isotopes in the target, $T_i(q^2)$ is the scattering matrix element at momentum transfer squared $q^2 = 2m_i\nu$, and f_i is the mass fraction of isotope i. A sum over final and average over initial polarizations is understood in $\overline{|T_i(q^2)|^2}$. The factor

$$\eta_i(q) = \int_{q/2\mu_i}^{\infty} \frac{f_\chi(v)}{v} \mathrm{d}^3 v, \tag{2}$$

with units of inverse velocity, incorporates the χ velocity distribution $f_\chi(v)$. For a Maxwellian distribution with velocity dispersion v , seen by an observer moving at speed v_O,

$$\eta_i(q) = \frac{1}{2v_O}\left[\mathrm{erf}\left(\frac{v_q + v_O}{\sqrt{2}v_{\mathrm{rms}}}\right) - \mathrm{erf}\left(\frac{v_q - v_O}{\sqrt{2}v_{\mathrm{rms}}}\right)\right], \tag{3}$$

with $v_q = q/2\mu_i$. For standard halo parameters, $\eta_i(q)$ is approximately exponential in the deposited energy ν. The previously-mentioned modulations enter the rate through $\eta_i(q)$.

The scattering matrix element $T(q^2)$ can be written as the Fourier transform

$$T(q^2) = \int \langle \mathrm{f}|V(\vec{r})|\mathrm{i}\rangle \mathrm{e}^{i\vec{q}\vec{r}/\hbar} \mathrm{d}\vec{r} \tag{4}$$

of a non-relativistic WIMP-nucleus potential

$$V(\vec{r}) = \sum_{\substack{\text{pointlike} \\ \text{nucleons} \\ n=\mathrm{p,n}}} (G_s^n + G_a^n \vec{\sigma}_\chi \vec{\sigma}_n)\, \delta(\vec{r} - \vec{r}_n). \tag{5}$$

The constants G_s^n and G_a^n are effective four-fermion coupling constants for nucleon-WIMP interactions, and are analogous to Fermi's constant G_F. G_s^n represents scalar[1] or spin-independent interactions, G_a^n axial[1] or spin-dependent interactions. Both terms are coherent in the quantum-mechanical sense when $qR_{\mathrm{nucleus}} \ll \hbar$, *i.e.* when the nucleus can be treated as pointlike and $T(q^2)$ can be taken as $T(0)$. At larger q, which can occur with heavy target nuclei, both terms are incoherent. Nuclear form factors $F(q^2)$, conventionally defined by $T(q^2) = T(0)F(q^2)$, should then be introduced. The scalar and spin form factors are in general different, reflecting the difference in the mass and spin distributions inside the nucleus.

The task of a theoretician is to provide a theoretical estimate of $T(q^2)$ starting from a particle-physics model. We accomplish this by stages, successively finding the WIMP-quark,

[1] Associated to scalar and axial vectors under 3d rotations.

the WIMP-nucleon and the WIMP-nucleus effective lagrangians. Step 1, finding the effective WIMP-quark lagrangian at small q^2, is analogous to going from the Standard Model to four-fermion interactions. Step 2 requires knowledge of the quark content of the nucleon, *i.e.* the contributions of different quarks to the nucleon mass and spin. Step 3 needs a nuclear model to describe how protons and neutrons are distributed in a nucleus.

This procedure is now illustrated for a Dirac neutrino and for a Majorana particle, an example of which is the neutralino.

Dirac neutrino

Step 1: a Dirac neutrino ν interacts with a quark q through the diagram in Fig. 1a. At $q^2 \ll m_Z^2$, the Z propagator reduces to $ig^{\mu\nu}/m_Z^2$, and the four-fermion amplitude reads

$$\sqrt{2}G_F\ \bar{\nu}(v_\nu - a_\nu\gamma_5)\gamma_\mu\nu\ \bar{\mathrm{q}}(v_\mathrm{q} - a_\mathrm{q}\gamma_5)\gamma^\mu\mathrm{q}, \tag{6}$$

with $v_\nu = a_\nu = \frac{1}{2}$, $a_\mathrm{q} = T_{3\mathrm{q}}$ and $v_\mathrm{q} = T_{3\mathrm{q}} - 2e_\mathrm{q}\sin^2\theta_W$. Here $\sin^2\theta_W \simeq 0.23$ and e_q and T_{3q} are the electric charge and the third component of the weak isospin of quark q. For a non-relativistic neutrino, only the time component of the vector current and the space components of the axial current survive. The first is spin-independent ($\bar{\nu}\gamma_0\nu \propto \nu^\dagger\nu$) and the second spin-dependent ($\bar{\nu}\vec{\gamma}\gamma_5\nu \propto \nu^\dagger\vec{\sigma}\nu$).

Step 2 for the vector part

$$\sqrt{2}G_F\, v_\nu v_\mathrm{q}\ \bar{\nu}\gamma_\mu\nu\ \bar{\mathrm{q}}\gamma^\mu\mathrm{q}, \tag{7}$$

because of vector current conservation, simply amounts to summing T_{3q} and e_q of the constituent quarks. For protons and neutrons one obtains respectively

$$G_s^\mathrm{p} = \frac{G_F}{\sqrt{2}}(1 - 4\sin^2\theta_W)v_\nu \tag{8}$$

$$C_s^\mathrm{n} = -\frac{G_F}{\sqrt{2}}v_\nu. \tag{9}$$

The interaction is mainly with the neutrons since $1 - 4\sin^2\theta_W \approx 0$.

Step 2 for the axial part

$$\sqrt{2}G_F\, a_\nu a_\mathrm{q}\ \bar{\nu}\gamma_\mu\gamma_5\nu\ \bar{\mathrm{q}}\gamma^\mu\gamma_5\mathrm{q}, \tag{10}$$

leads to the four-fermion coupling constants

$$G_a^\mathrm{p} = \sqrt{2}G_F a_\nu\left(a_\mathrm{u}\Delta\mathrm{u} + a_\mathrm{d}\Delta\mathrm{d} + a_\mathrm{s}\Delta\mathrm{s}\right), \tag{11}$$

$$C_a^\mathrm{n} = \sqrt{2}G_F a_\nu\left(a_\mathrm{u}\Delta\mathrm{d} + a_\mathrm{d}\Delta\mathrm{u} + a_\mathrm{s}\Delta\mathrm{s}\right). \tag{12}$$

Here Δq is the fraction of the proton spin carried by quark q, $\frac{1}{2}\langle p|\bar{q}\gamma_\mu\gamma_5 q|p\rangle = \Delta q s_\mu$. It can be obtained[9] from data on neutron and hyperon β-decay, which give $\Delta u - \Delta d = 1.2573 \pm 0.0028$ and $\Delta u + \Delta d - 2\Delta s = 0.59 \pm 0.03$, respectively. The contribution of the strange quark is $\Delta s = 0$ in the naive quark model, $\Delta s = -0.11 \pm 0.03 \pm \cdots$ from deep inelastic data, and $\Delta s = -0.15 \pm 0.09$ from elastic $\nu p \to \nu p$ data.

Step 3 for the spin-independent part introduces the nuclear mass form factor $F_{\rm mass}(q^2)$, and results in

$$\overline{\left|T_s(q^2)\right|^2} = \left|ZG_s^{\rm p} + NG_s^{\rm n}\right|^2 \left|F_{\rm mass}(q^2)\right|^2, \tag{13}$$

where N (Z) is the number of neutrons (protons) in the nucleus. Neutron scattering off nuclei suggests that $F_{\rm mass}(q^2) \simeq F_{\rm e.m.}(q^2)$, the electromagnetic form factor. The electric charge distribution is well-described by a Fermi or Woods-Saxon form,[10] whose Fourier transform is indistinguishable from the convenient analytic expression[11]

$$F_{\rm mass}(q^2) \simeq \frac{3j_1(qR)}{qR} e^{-\frac{1}{2}(qs)^2}. \tag{14}$$

The electromagnetic radius R and the surface thickness s can be obtained by fitting electron scattering data,[10] or can be roughly approximated by $R \approx A^{1/3}$ fm and $s \approx 1$ fm.[4] $F_{\rm e\,m.}(q^2)$ presents diffraction zeros when the modified Bessel function $j_1(qR) = 0$, the first of which occurs at $qR \simeq 4.2$. In electron scattering, these diffraction zeros are filled in, because due to the long-range Coulomb attraction the electron wave function is distorted from a simple plane wave and the form factor is not simply the Fourier transform of the charge density. The short-range nature of WIMP-nucleus interactions make us expect no wave function distortion, and diffraction zeros remain.[2] The first diffraction zero is important in assessing bounds from some present-day detectors.[12]

Step 3 for the spin-dependent part requires the expectation values of the total spin of protons $\langle S_{\rm p}\rangle$ and neutrons $\langle S_{\rm n}\rangle$ separately. At $q = 0$,

$$\overline{\left|T_a(0)\right|^2} = \frac{4(J+1)}{J}\left|G_a^{\rm p}\langle S_{\rm p}\rangle + G_a^{\rm n}\langle S_{\rm n}\rangle\right|^2, \tag{15}$$

where J is the nuclear spin. Even-even nuclei, with even numbers of protons and of neutrons, do not have spin, and for them $T_a(0) = 0$. For even-odd nuclei with $J \neq 0$, a nuclear model is needed to estimate $\langle S_{\rm p}\rangle$ and $\langle S_{\rm n}\rangle$. For instance, ^{73}Ge is an odd-neutron nucleus with $J = \frac{9}{2}$. The single-particle shell model[1, 13] gives

$$\langle S_{\rm n}\rangle = \frac{1}{2}\left[1 + \frac{\frac{3}{4} - l(l+1)}{j(j+1)}\right] = 0.50, \qquad \langle S_{\rm p}\rangle = 0; \tag{16}$$

[2]The first corrections are at a level of 10^{-6} and come from neglected higher powers of the incoming WIMP velocity.

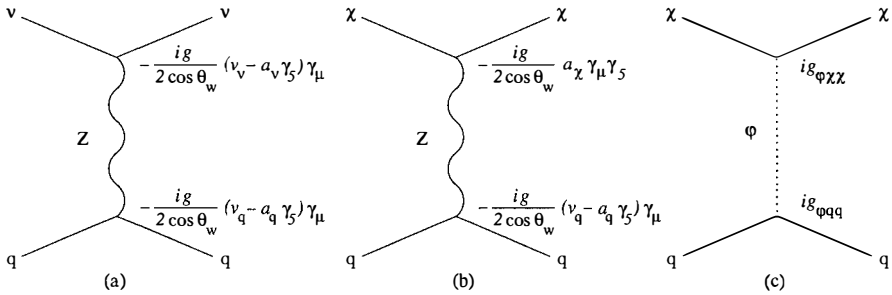


Figure 1: *Examples of WIMP-quark scattering.*

the odd-group model,[14] in which the odd-nucleon spin is related to the nuclear magnetic moment μ and gyromagnetic factors $g^{L,S}_{\rm n,p}$, gives

$$\langle S_{\rm n}\rangle = \frac{\mu - g^L_{\rm n} J}{g^S_{\rm n} - g^L_{\rm n}} = 0.23, \qquad \langle S_{\rm p}\rangle = 0; \tag{17}$$

a more sophisticated interacting shell model[15] gives

$$\langle S_{\rm n}\rangle = 0.468, \qquad \langle S_{\rm p}\rangle = 0.011. \tag{18}$$

The proton might have a small but non-zero contribution to the cross section, which might change the relative merits of different nuclei for dark matter searches.

At $q \neq 0$, nuclear spin form factors are needed. The neutron and proton contributions differ, and at present only complex calculations[15, 16] for specific nuclei provide an estimate of the isoscalar and isovector spin form factors $F^0_{\rm spin}(q^2)$ and $F^1_{\rm spin}(q^2)$, in terms of which

$$\overline{\left|T_a(q^2)\right|^2} = \frac{J+1}{J}\left|(G^{\rm p}_a + G^{\rm n}_a)\langle S_{\rm p} + S_{\rm n}\rangle F^0_{\rm spin}(q^2) + (G^{\rm p}_a - G^{\rm n}_a)\langle S_{\rm p} - S_{\rm n}\rangle F^1_{\rm spin}(q^2)\right|^2, \tag{19}$$

The results of these calculations can be conveniently resumed by the approximate expressions

$$F^0_{\rm spin}(q^2) \simeq \exp\left(-\frac{r_0^2 q^2}{\hbar^2}\right), \qquad F^1_{\rm spin}(q^2) \simeq \exp\left(-\frac{r_1^2 q^2}{\hbar^2} + i\frac{cq}{\hbar}\right), \tag{20}$$

with parameters given in the following table for selected nuclei:

	J	$\nu_{\rm max}$/keV	$\langle S_{\rm p}\rangle$	$\langle S_{\rm n}\rangle$	r_0/fm	r_1/fm	c/fm	valid for ν/keV<
^{73}Ge	$\frac{9}{2}$	540	0.011	0.468	1.971	2.146	-0.246	55
^{28}Si	$\frac{1}{2}$	216	-0.0019	0.133	1.302	1.548	-0.320	145
^{27}A	$\frac{5}{2}$	100	0.3430	0.269	1.378	1.600	0.196	$\nu_{\rm max}$
^{39}K	$\frac{3}{2}$	145	-0.184	0.054	1.746	1.847	0.371	$\nu_{\rm max}$

Majorana fermion

A Majorana fermion is a spin-$\frac{1}{2}$ particle that coincides with its antiparticle. It carries no conserved quantum number. It has neither vector nor tensor currents. Of the remaining pseudoscalar, scalar and axial currents, only the last two have a non-vanishing non-relativistic limit, spin-independent the first ($\bar{\chi}\chi \propto \chi^\dagger\chi$) and spin-dependent the second ($\bar{\chi}\vec{\gamma}\gamma_5\chi \propto \chi^\dagger\vec{\sigma}\chi$).

Axial currents may arise from exchange of a Z boson as in fig. 1b, and the analysis is then analogous to that in the previous section, with the obvious replacement of a_ν with a_χ.

Scalar currents originate from exchange of a scalar particle φ, *e.g.* as in Fig. 1c. At small q^2, the φ propagator reduces to $-i/m_\varphi^2$ and the four-fermion amplitude reads

$$-\frac{g_{\varphi\chi\chi}g_{\varphi\mathrm{qq}}}{m_\varphi^2}\,\bar{\chi}\chi\,\bar{\mathrm{q}}\mathrm{q}. \tag{21}$$

For a nucleon $n = \mathrm{p}, \mathrm{n}$ one then obtains

$$G_s^n = -\frac{g_{\varphi\chi\chi}}{m_\varphi^2}\sum_{\mathrm{q}} g_{\varphi\mathrm{qq}}\langle n|\bar{\mathrm{q}}\mathrm{q}|n\rangle. \tag{22}$$

For example, in the case of the neutralino with exchange of the lightest supersymmetric Higgs boson, the sum over quarks is explicitly

$$\frac{g}{2m_\mathrm{W}}\left[\frac{\cos\alpha}{\sin\beta}\langle m_\mathrm{u}\bar{\mathrm{u}}\mathrm{u} + m_\mathrm{c}\bar{\mathrm{c}}\mathrm{c} + m_\mathrm{t}\bar{\mathrm{t}}\mathrm{t}\rangle - \frac{\sin\alpha}{\cos\beta}\langle m_\mathrm{d}\bar{\mathrm{d}}\mathrm{d} + m_\mathrm{s}\bar{\mathrm{s}}\mathrm{s} + m_\mathrm{b}\bar{\mathrm{b}}\mathrm{b}\rangle\right]. \tag{23}$$

The scalar quark content of the nucleon $\langle n|\bar{\mathrm{q}}\mathrm{q}|n\rangle$ can be extracted from data with the help of chiral perturbation theory, π-nucleon scattering and heavy quark expansion.[17] The result is

$$\langle m_\mathrm{u}\bar{\mathrm{u}}\mathrm{u}\rangle \simeq \langle m_\mathrm{d}\bar{\mathrm{d}}\mathrm{d}\rangle \simeq 30\,\mathrm{MeV}/c^2, \qquad \langle m_\mathrm{s}\bar{\mathrm{s}}\mathrm{s}\rangle \simeq 60\text{–}120\,\mathrm{MeV}/c^2, \tag{24}$$

$$\langle m_\mathrm{c}\bar{\mathrm{c}}\mathrm{c}\rangle = \langle m_\mathrm{b}\bar{\mathrm{b}}\mathrm{b}\rangle = \langle m_\mathrm{t}\bar{\mathrm{t}}\mathrm{t}\rangle = \frac{2}{27}\left(m_\mathrm{p} - \sum_{\mathrm{q}=\mathrm{u},\mathrm{d},\mathrm{s}}\langle m_\mathrm{q}\bar{\mathrm{q}}\mathrm{q}\rangle\right) \simeq 60\,\mathrm{MeV}/c^2. \tag{25}$$

The strange quark contribution is uncertain by a factor of 2. Step 3 is analogous to the Dirac neutrino case, and leads to eq. (13) with four-fermion couplings given by (22).

Neutralino

Supersymmetry and the neutralino have been presented by Jungman at this conference. The neutralino has both spin-dependent and spin-independent interactions with nuclei, the former mediated by Z boson and squarks, the latter by Higgs bosons and squarks. The general formalism of the preceding sections can be used. In the limit of heavy squarks $\tilde{\mathrm{q}}_k$, the effective four-fermion constants are given by

$$C_s^\mathrm{p} \simeq G_s^\mathrm{n} = \sum_{\mathrm{q}=\mathrm{u},\mathrm{d},\mathrm{s},\mathrm{c},\mathrm{b},\mathrm{t}}\langle\bar{\mathrm{q}}\mathrm{q}\rangle\left(-\sum_{h=\mathbf{H_1},\mathbf{H_2}}\frac{g_{h\chi\chi}g_{h\mathrm{qq}}}{m_h^2} + \frac{1}{2}\sum_{k=1}^{6}\frac{g_{L\tilde{\mathrm{q}}_k\chi\mathrm{q}}g_{R\tilde{\mathrm{q}}_k\chi\mathrm{q}}}{m_{\tilde{\mathrm{q}}_k}^2}\right), \tag{26}$$

$$G_a^p = \sum_{q=u,d,s} \Delta q \left(\frac{g_{Z\chi\chi} g_{Zqq}}{m_Z^2} + \frac{1}{8} \sum_{k=1}^{6} \frac{g^2_{\tilde{L}_k \chi q} + g^2_{\tilde{R}_k \chi q}}{m^2_{\tilde{q}_k}} \right), \qquad G_a^n = G_a^p(\Delta u \leftrightarrow \Delta d). \tag{27}$$

Expressions for the elementary vertices g_{ijk} can be found in ref. 18.

Predictions in supersymmetric models suffer from the presence of many unknown parameters. Two extreme attitudes are a phenomenological approach in which what is not excluded is allowed, and a grand-unified approach in which coupling constants and masses are unified at some high energy scale. Fig. 2 shows examples of calculated event rates in ^{76}Ge, each point representing a choice of model parameters: "predictions" may well span 10 orders of magnitude in a phenomenological approach[18] and 2 orders of magnitude in a more restricted scenario.[19]

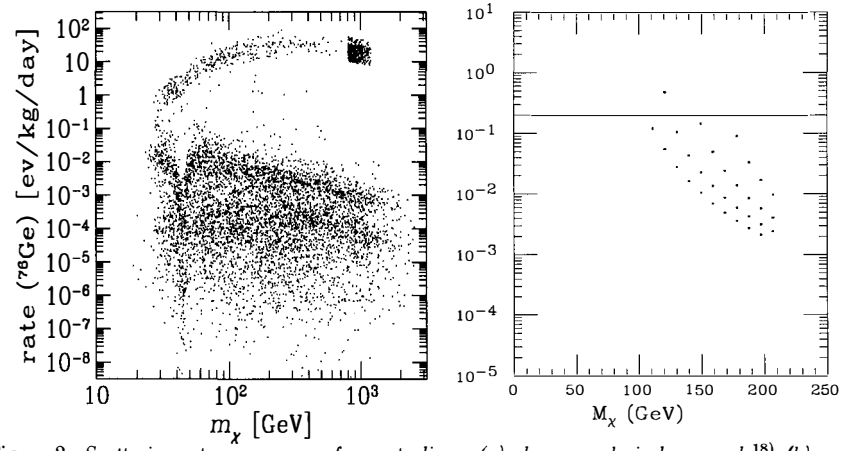


Figure 2: *Scattering rate versus mass for neutralinos: (a) phenomenological approach,*[18] *(b) grand-unified approach.*[19]

Underabundant dark matter relics

Given a particle-physics model, the relic density of a species, a WIMP χ in particular, is a calculable and definite quantity. Often it happens that the computed relic density Ω_χ is (much) smaller than the dark matter density in the Universe. For this reason, some authors simply neglect this case. But even if these WIMPs constitute only a fraction of the dark matter, they generally have quite high scattering cross sections off nuclei, because of an approximate inverse proportionality of the χ relic density and the χ–nucleus cross section. However, the scattering rate also includes the χ *halo* density ρ_χ. It is reasonable that ρ_χ is only a fraction of the local dark matter density ρ_{DM}, but which precise fraction it is depends on the model for

galaxy formation. If both the main and the χ components of dark matter are cold, we expect them to behave similarly under gravitation, so that the halo fraction f_χ might be equal to the universal fraction $\Omega_\chi/\Omega_{\rm DM}$. Unfortunately, $\Omega_{\rm DM}$ is poorly known: it can range from ≈ 0.01 for dark matter associated with galactic halos to ≈ 1 for a smooth universal component. In fig. 3, the suppression of scattering rates due to rescaling of the neutralino halo density by a universal fraction with $\Omega_{\rm DM}h^2 = 0.025$ is apparent to the left of the dashed line. This suppression must be included for consistency when setting bounds on particle-physics models.

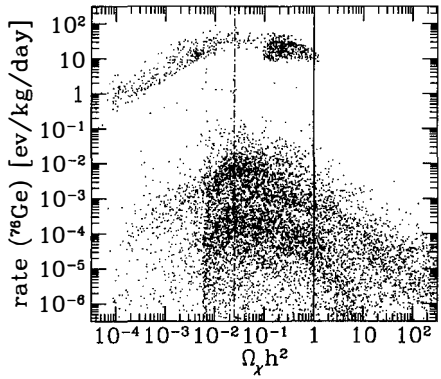


Figure 3: *Scattering rate versus relic density for neutralinos (from ref. 18).*

References

[1] M.W. Goodman and E. Witten, Phys. Rev. D31 (1986) 3059.

[2] J.R. Primack, B. Sadoulet and D. Seckel, Annu. Rev. Nucl. Sci. 38 (1988) 751.

[3] P.F. Smith and J.D. Lewin, Phys. Rep. 187 (1990) 203.

[4] J. Engel, S. Pittel and P. Vogel, Int. J. Mod. Phys. E1 (1992) 1.

[5] J.A.R. Caldwell and J.P. Ostriker, Ap. J. 251 (1981) 61.

[6] A. Drukier, K. Freese and D. Spergel, Phys. Rev. D33 (1986) 3495.

[7] J.I. Collar and F.T. Avignone, Phys. Rev. D47 (1993) 5238.

[8] D.N. Spergel, Phys. Rev. D37 (1988) 1353.

[9] A recent analysis can be found in J. Ellis and M. Karliner, hep-ph/9601280.

[10] R. Hofstadter, Rev. Mod. Phys. 28 (1956) 214; "Electron scattering and nuclear and nucleon structure" (Benjamin, 1963).

[11] R.H. Helm, Phys. Rev. 104 (1956) 1466.

[12] A. Bottino *et al.*, Phys. Lett. B295 (1992) 330.

[13] J. Ellis and R. Flores, Nucl. Phys. B307 (1988) 883; Phys. Lett. B263 (1991) 259.

[14] J. Engel and P. Vogel, Phys. Rev. D40 (1989) 3132.

[15] M. Ted Ressell *et al.*, Phys. Rev. D48 (1993) 5519.

[16] J. Engel, Phys. Lett. B264 (1991) 114; J. Engel *et al.*, Phys. Rev. C52 (1995) 2216.

[17] M.A. Shifman, A.I. Vainshtein and V.I. Zakharov, Phys. Lett. B78 (1978) 443; T.P. Cheng, Phys. Rev. D38 (1988) 2869; H.Y. Cheng, Phys. Lett. B219 (1989) 347; J. Gasser, H. Leutweiler and M.E. Sainio, Phys. Lett. B253 (1991) 252.

[18] L. Bergström and P. Gondolo, hep-ph/9510252, to appear in Astropart. Phys.

[19] V. Berezinsky *et al.*, hep-ph/9508249.

Résumé

La matière noire qui forme le halo de notre galaxie pourrait être constituée de particules interagissant faiblement avec la matière ordinaire : les WIMPs. Malgré le faible taux d'interaction attendu entre ces particules et le cristal d'un détecteur, la détection directe de WIMPs apparaît possible et plusieurs expériences sont en cours actuellement. Nous décrivons, dans cet article, les éléments de base permettant de calculer ces taux d'interaction, en insistant sur les principales sources d'incertitudes théoriques.

Introduction and Summary talks for the Session :

"PARTICLE DARK MATTER DIRECT DETECTION"

by L. MOSCA (C.E. - SACLAY/DAPNIA/SPP)

Direct Particle Dark Matter detection experiments are briefly reviewed by emphasizing the main problems encountered in these approaches and the possible strategies to overcome them.

1. INTRODUCTION [1]

In this paper I will mostly refer to **WIMPs** (Weakly Interacting Massive Particles) candidates. I will consider the principle of **direct detection,** the main detectors involved, the problems to be solved and the consequent possible strategies to be adopted, the present status of the art both in terms of R&D and physical results (see for more details the individual contributions to this session).

Particle (non-baryonic) candidates can be **non-thermal** relics (like Axions or GUT-Monopols) or **thermal** relics. These last, in turn, can belong to the "**Hot Dark Matter**" or to the "**Cold Dark Matter**" depending if, at the moment of their decoupling they were relativistic or not relativistic.

The **Hot Dark Matter** can be made of **light neutrinos** (ν_e, ν_μ, ν_τ), while **Cold Dark Matter** can be made of **WIMPs** (that is **massive neutrinos**, of both Dirac or Majorana type, or **Supersymmetric - SUSY - particles**) or **SIMPs** (Strongly Interacting Massive Particles).

Masses and **cross sections** of WIMPs, interacting with ordinary matter, are **largely unknown**. Masses can range between 1 GeV and 1 TeV, while cross sections on nuclei can range within some orders of magnitude around a typical "weak interaction value" ($\approx 10^{-38}$ cm^2).

Of special interest is the **Lightest** of the Supersymmetric Particles (**LSP**), the **neutralino**. In the simplest versions of SUSY (with R-parity conservation) the neutralino is **stable**, so that it is a good candidate as a **relic** particle in the Universe.

The **neutralino** interaction with ordinary matter is depending on several parameters (≥ 5) which are free, to some extent, in the present formulations of the SUSY theory [2]. These parameters control the actual relative proportion of **two basic couplings** :

α) SPIN-INDEPENDENT (COHERENT) : $\sigma_{WIMPS\text{-}N} \approx A^2$, where A is the Atomic mass number of the target nucleus N, and

β) SPIN - DEPENDENT (AXIAL) : $\sigma_{WIMPS\text{-}N} \approx C\ \lambda^2\ J\ (J+1)$, where C = < Hadronic Matrix Element > (squared) of the "odd nucleon" in the nucleus N, and λ is the Landé factor of the "single-particle" (or of the "odd-group") Shell Model of nuclei.

2. DIRECT DETECTION : *PRINCIPLE AND PROBLEMS*

WIMPS can be detected **directly** by elastic scattering on nuclei [3], that is by measuring the recoil energy of the scattered nucleus. In spite of the simplicity of this detection principle **3 kinds of difficulties** appear immediately :

a) the **rates** of these events are **small** and depend on the type of dominant coupling. Typical values are : 1 to 100 events/day/kg for a dominant **coherent coupling** (neutralino with both zino and Higgsino(s) components, Dirac ν, and s-neutrino), and 0.01 to 1 events/day/kg for, a dominant **spin dependent coupling** (neutralino which is mainly photino and/or zino or is pure Higgsino, and Majorana ν).

b) the **recoil energies** are also **small** (typically from $\approx$ 1 to $\approx$ 30 keV)

c) and, last but not least, the **background** is **comparatively very high**. This is of two origins :

α) the secondary **cosmic rays** which are disturbing in two ways : by directly depositing energy in the detectors and by activating the nuclei of the materials composing or surrounding these detectors (cosmogenic radioactivity).

For this reason it is essential to install the detectors **deep underground**.

β) the **radioactive background** . For that it is necessary :

- to **shield** (with Pb, Cu, etc) **the detectors** from the environment
- to use **very low activity materials** (by selection and/or purification)
- to maximise the **intrinsic rejection power of the detectors** against low energy electrons, by Pulse Shape Discrimination (PSD) of the signals (scintillators) and/or by the simultaneous detection of 2 different physical quantities : heat and ionisation or heat and scintillation (bolometers).

Now **the important question** is : which are the possible **signatures** for the WIMP direct detection ?

Due to the orbital motion of the Earth around the Sun, the **average WIMP velocity** ($\beta \approx 10^{-3}$), relative to the Earth, is **not constant** in time, thus producing a **modulation,** between June and December, **of the average WIMP flux** (and of the average nuclear recoil energy) of the order of 5 to 10%.

This **annual modulation** potentially represents the **basic signature of WIMPs** in their **direct detection** , by proving in this way their **astrophysical origin**.

It is **important,** in this respect, **to realise** :

- **how large must be the mass of a detector** in order **to be sensitive to such an effect** : a mass $_{det}$ (kg) $\approx$ 50 / WIMPs interaction Rate (evts/kg/day) is required for a 5 Standard Deviations evidence of a signal.
- that the **intrinsic background rejection power must be**, as much as possible, of the type "**event by event**" to be useful for this purpose.

The **diurnal modulation,** based on the WIMPs interactions inside the Earth, is also a **possible and interesting signature** for sufficiently high cross sections [7c].

An **additional signature** can be obtained by considering that the event rate is depending on the type of nuclei used as targets, in a way predicted by the theory.

3. EXPERIMENTAL REVIEW

Let us now review the different kinds of WIMPS direct detection experiments which depend on the different types of detector used or under development.

We consider first the **classical semiconductor detectors** (Germanium and Silicium), at liquid N_2 temperature, then the **crystal and liquid scintillators,** at room temperature, the **Bolometers**, at few tens of millikelvin, and finally the **Track Projection Chambers** (TPC) and the **Superconducting grains**.

i) "Classical" Semiconductor Ge (Si) detectors :

Traditionally used for **2β experiments** (**MeV energy range**), these detectors have been **optimised also for Dark Matter (WIMPs) search** (**energy range of a few keV**).

Table 1 summarizes the experimental situation from this point of view : here the good performance of the Heidelberg-Moscow experiment[8] have to be stressed, although the threshold is much higher than in the other quoted experiments.

Limitations of these detectors are the **microphonics problem at low energies**, the **electronic noise** and, more generally, the **technical and financial problems** rapidly increasing with the mass of the crystals.

Table 1. Semiconductor Experiments

Detec--tor crystal	Collaboration	Underground Laboratory	Mass of the crystal	Threshold (keV) (+)	Rate (*) evts/keV /kg/day)	Comments (year)
Ge	Ahlen/ Avignone/ /Drukier [4]	Homestake (≈4400 mwe)	0.25 kg to 1 kg	4	≈ 2	(1987-1991)
Ge	UCSB/UCB/ /LBL [5]	Oroville (≈500 mwe)	900 g	3	≈ 3	(1988)
Ge	Caltech/Psi/ /Neuchatel [6]	St. Gotthard (≈3000 mwe)	800 g	2	≈ 3	low threshold (1991)
Ge	Zaragoza/PNL/ /USC [7a]	Canfranc tunnel (675 mwe)	234 g	3	≈ 8	Cosme - 2 detector (1993)
Ge	Heidelberg/ /Moscow [8]	Gran Sasso (≈3500 mwe)	2.9 kg	12	≈ 0.2	low rate (1994) high threshold
Si	UCSB/UCB/ /LBL/SACLAY [9]	Oroville (≈500 mwe)	4 x17g	1.1	≈ 60	for Cosmions (high σ and low masses) (1990)

(+) Electron Equivalent Energy (*) near threshold, after "microphonic events" subtraction

ii) Scintillation detectors :

In spite of their **poorer resolution** compared to Ge detectors, **scintillators play** at present **an important role** in this field for several reasons. Scintillation detection is a **well known and simple technique** (usually no cooling needed), the ratio **cost/mass** is **reasonably low,** and there is a **wide choice of target nuclei.**

In particular some of the scintillators (e.g. **NaI(Tl)**) have **a good light output** and **a demonstrated** (although problematic at low energy) Pulse Shape Discrimination (**PSD**). In addition scintillators include **non zero spin target nuclei** (^{23}Na, ^{127}I, ^{19}F, ^{129}Xe,.....) . Other specific reasons of interest are the **very high predicted axial coupling for the ^{19}F nucleus** and the **high** estimated **Quenching Factor of liquid Xenon** (see below).

The **Quenching Factor (QF)** for nuclear recoils (the scintillation efficiency for nuclear recoils relative to that of electrons of the same kinetic energy) is measured by using a **neutron beam** of a few MeVof kinetic energy. The results are quoted on Table2

Table 2. "Quenching Factor" of nuclear recoils, measured with neutron beams.

Scintillator	Target Nucleus	Saclay and IN2P3[10] at Bruyères Le Chatel	Imp. College/ Oxford/Rutherford [11]	Osaka (Japan) [12]
NaI(Tl)	^{23}Na	0.25± 0.02	≈ 0.30	0.4 ± 0.2
	^{127}I	0.08 ± 0.01	≈ 0.08	0.05 ± 0.02
CaF2 (Eu)	^{40}Ca	0.049 ± 0.005	0.08 ± 0.01	-------
	^{19}F	0.069 ± 0.005	0.12 ± 0.01	-------

- for Liquid Xenon : QF ≈ 0.8 (as estimated from Lindhard theory + recombination effects (while for α particles the measured QF ≈ 1.2) [13]

Moreover let us remind that :

- for Semiconductors : QF(Ge) ≈ 0.25 [14]; QF(Si) ≈ 0.30 [15]
- for Bolometers (Heat channel) : QF ≈ 1.0 (to be checked at low energy)

The present situation of the scintillator experiments is summarized in **Table 3.**

Table 3. Scintillator Experiments

Collaboration	Underground Laboratories	Type of scintil-lators	Mass (kg)	Thresh (keV) (+)	Rate near threshold (evts/keV /kg/day)	Resolu-tion σ/E at 60keV	Comments
BPRS (Beijing, Paris, Roma, Saclay) [10,17,18]	GranSasso/Fréjus Mentogou	NaI(Tl) NaI(Tl) CaF_2(Eu)	0.76 7 0.37	≈ 4 ≈ 4 ≈ 4	≈ 7 ≈ 2 ≈15	≈ 7 % ≈ 7 % ≈12 %	dedicated expt.(R&D up to 10 kg)
Saragoza [7b]	Canfranc tunnel	NaI(Tl)	3×10.7	≈ 8	≈ 8 (at 18 keV)	≈12 %	-------
Osaka (Japan) [12, 16]	Kamioka mine (≈2700 mwe) NARA (Oto) (≈500 mwe)	NaI(Tl) CaF2(Eu)	(17x) 36.5 25 x 0.29	≈ 5	≈ 6	≈10 %	elastic and inelastic. preliminary
Imp.College/ /Oxford/ [11] /Rutherford	Boulby mine (≈3000 mwe)	NaI(Tl)	1 6	≈ 4 ≈ 4	≈ 7 ≈ 2	≈10 %	dedicated expt.
DAMA (Roma) [13]	Gran Sasso	Liq. Xe	6.5	≈ 10	≈ 3	≈26 %	dedicated expt.

(+) Electron Equivalent Energy

iii) Bolometers :

These are **very sensitive thermometers** (**including**, by extension, the associated **heat absorber**) operated at very low temperature (≈ 10 mK).

Typical absorbers are Sapphire(Al_2O_3), Germanium, LiF, TeO_2, etc (see **Table 4**)

Why using bolometers to detect WIMPs ?

As discussed in section 2, **bolometers allow**, at present, **a simultaneous measurement of thermal information** (lost in other detectors) **and ionization** (or **scintillation**) thus allowing particles vs nuclear recoils discrimination at low energy (background rejection). Moreover, the **quanta of thermal energy** (phonons) are **very tiny** (10^{-4} - 10^{-5} **eV**) so a potentially **very low threshold** and **very good resolution** are expected.

Table 4. Bolometers : present status of the R&D and data taking.

Colla- -boration	Undergr. Laborat.	Crystal	Mass (g)	Thresh (keV) (*)	Temp (mK)	Rate evts/keV /kg/day	Resolution σ/E	Com- -ments
CfPA(Berkeley Stanford/UCSB /INR) Sadoulet [19] CDMS	shallow site at Stanford 17 mwe	Ge	62	≈ 3	24	____	0.5% at 60keV (1% Ionizat.)	both : - HEAT - IONIZ.
MUNICH Max Plank/Gar- -ching/Oxford [20] CRESST	Gran Sasso (1995 - 96)	Al_2O_3 (Sapph- -ire)	32 262	≈ 0.3	15 44	____	3% at 1.5 keV 1.5% at 6 keV ≈1 % at 1MeV 4.5% at 60 keV	only HEAT (WIMPS < 15 GeV)
CEA/IN_2P_3/ INSU (France) [21] EDELWEISS	Fréjus 4800 mwe	Al_2O_3 Ge	24 7 (from Berkeley)	≈ 2.0 ≈1.7	55 34	≈25(2nd- -run) at 16 keV	2.5% at 60 keV 3.0 % (2.0% Ionizat.)	only HEAT both:HEAT and IONIZ.
TOKYO [22]	Kamioka	LiF	2.8	≈ 4	12		6% at 60 keV	only HEAT
Milano Univ. Fiorini et al. [23]	Gran Sasso 3400 mwe	TeO_2 " LiF NaF $CdWO_4$ CaF_2	340 73 100 30 60 2	___	≈ 25	≈ 9 at13 keV _____	0.9% at100 keV 1 % at 60 keV 5 % at 1 MeV 2 % at " 0.2 % at " 2 % at "	for ββ+DM " forDarkMat for DarkMat HEAT+SC? HEAT+SC.

(*) Real Energy (Heat channel) (provided that the QF ≈ 1.0 down to low energies).

The **problem now** is to minimise **microphonics, electronic noise, radioactive background at low energy** and reach long time (years) **stable operating conditions**

with large masses of crystals (at least hundred kg) **not only** in the heat detection, **but also for the ionization** (or scintillation) detection.

The **present status** of the R&D and of data taking is summarized in **Table 4**.

Concerning heat detection alone, impressive results (threshold and resolution at very low energy) have been obtained by the Munich group [20].

A preliminary but **very encouraging result**, obtained with a massive (340 g) **TeO_2 crystal,** has been presented by the Milano-Gran Sasso group [23c and Fig1]: improvements are expected, in particular, from an ancient Roman lead shielding in preparation.

Concerning the **simultaneous measurement of heat and ionization,** the first successful result has been obtained by the **Berkeley group** [19], while the **French group** [21] has recently well improved the stability of the ionization channel (a crucial point of their detectors!). On the other hand the **Milano**[24], **Roma**[25], **Lyon**(IPN)/**Paris** (ENS)[26a] and **Lyon**(IPN)/**Orsay** (IAS) [26b] **groups** are developing **scintillation detection at low temperatures**, although, for the moment, at a much higher threshold ($\geq$ 100 keV).

iv) Track Projection Chambers (TPC) :

A **TPC** detector **at low pressure** has been first investigated by a Saclay group [27] (Ethane) and then by a San Diego group [28] (Hydrogen). The **main advantage** here is the **directionality** detection possibility, and the **main drawback** is the **very low density** of the detector medium.

v) Superconducting grains :

Two experiments are in preparation :

Orpheus (Bern, Villigen, Annecy/LAPP, Clausthal-Zellerfeld Collaboration) with a **13 g Sn** test-detector **at 10 mk** [29], and a **Lisbon, Zaragoza, Paris** Collaboration [30] with a pilot experiment using **100 g Sn superconducting grains,** at **100 mK**, in the **Canfranc** Laboratory. They claim an energy threshold < 2 keV and a background rejection factor of 95%.

vi) Ancient mica :

A very original project considers that **mica** can have stored, for $\approx$ 1 Gy, tracks produced by nuclear recoils from WIMPs. The **fission neutrons** background is a major challenge to the potential WIMP **signature** from **tracks directionality** [31].

From experimental rates to exclusion plots

By **comparing** the **measured** energy spectra **with** the **calculated** WIMPs **nuclear recoil spectra**, and considering that the signal cannot be bigger than the measured spectrum, it is possible to **deduce excluded regions in the** cross section versus mass plane ("**exclusion plots**") for WIMPs **interacting** with **each** type of **nuclei** in a given crystal detector.

The **ingredients** for that are: **quenching factors** [see section 3ii)], **astrophysical assumptions** on the WIMP halo **velocity distribution** and WIMP **local density** ($\approx$ 0.3 GeV/ cm^3, multiplied by an appropriated "**rescaling factor**") and **nuclear form factors**, relevant for heavy nuclei (typically A > 50), these last depending on the type of coupling : coherent or spin-dependent.

Moreover, to allow a **comparison between** results coming from **different target nuclei** it is necessary to obtain , e.g., σ (WIMP - proton) from σ (WIMP - Nucleus) and collect on the same diagram : σ (WIMP - proton) vs WIMP mass, the curves obtained for the different experiments.

Figs 1 and 2 summarize the **main existing results** for coherent and spin-dependent (axial) couplings respectively. The exclusion-plots are obtained by using the **same analysis program** applied to the energy spectra of the different experiments, published or presented at recent conferences.

Fig 1. Exclusion plots (Coherent coupling)

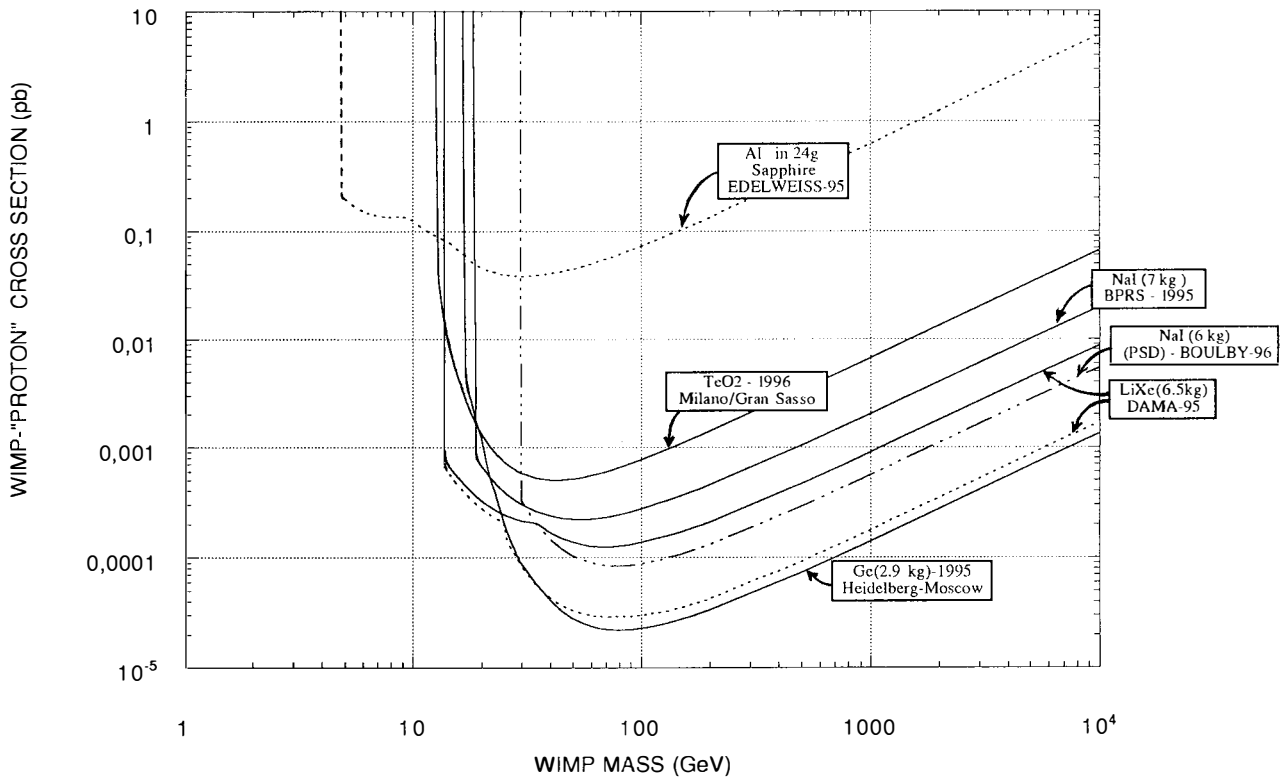
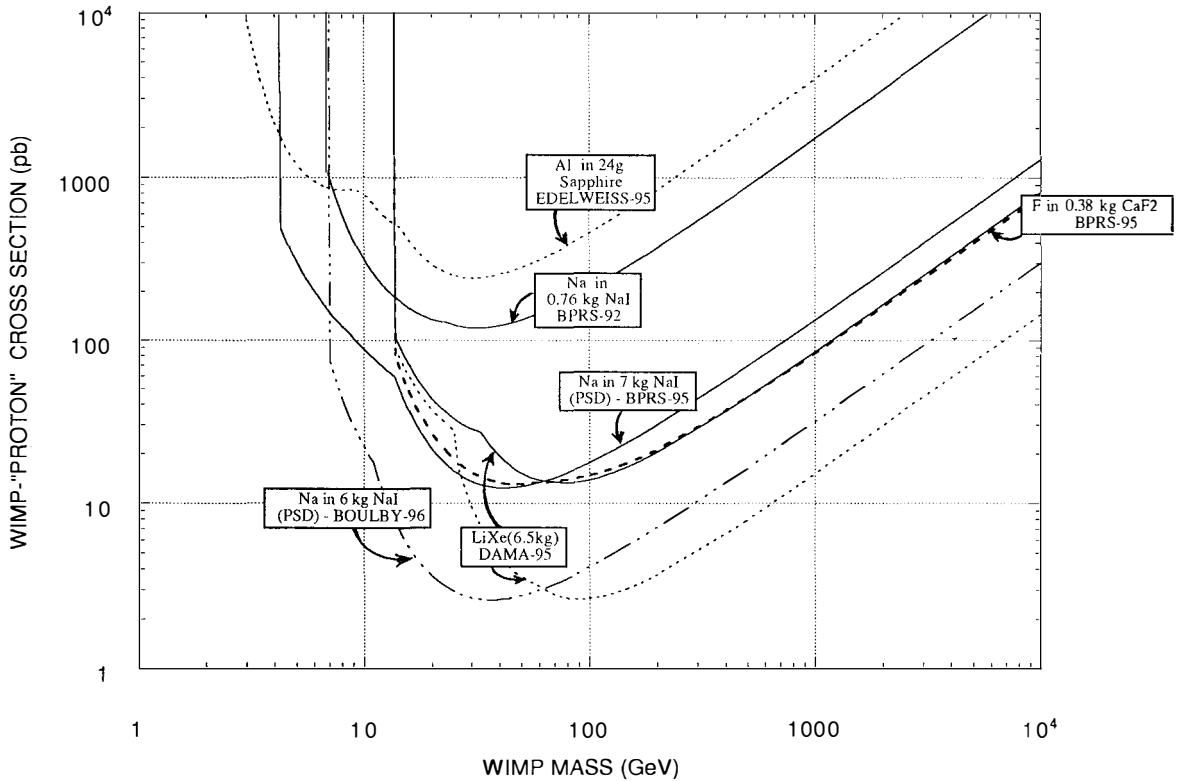


Fig 2. Exclusion plots (Axial coupling)



The basic assumptions in this calculation "exercise" are :

i) a **spherical** and **isotropic** WIMPs **halo** with a **Maxwellian velocity distribution** around the average value of 270 km/sec. The **speed of the solar system** with respect to this "static" halo is taken equal to 220 km/sec (the annual modulation effect is neglected at this stage). The WIMPs **escape velocity** in the galactic frame is taken equal to 800 km/sec.

ii) an $\mathbf{A^2}$ dependence of the **coherent cross section** (in view of the **neutralino** candidate : see the Introduction).

iii) **spin factors** $[\lambda^2 J(J+1)]$ and **hadronic marix elements** on "odd proton" or "odd neutron" for the different nuclei taken from Ref.[32].

iv) **Form Factors** for coherent coupling from Ref.[33] and for spin-dependent coupling from Ref.[32].

Concerning the individual experiments :

a) for the Boulby mine (NaI of 6 kg) and BPRS (NaI of 7 kg) experiments the results of a preliminary PSD analysis have been used [11].

b) for the Liquid-Xenon experiment two curves are quoted for each type of coupling: dotted lines are obtained from an energy spectrum with Gas-Xenon data subtracted (a delicate exercise mainly due to the different amount of energy deposition in both conditions), while full lines are obtained from the non-subtracted energy spectrum [13].

4. OUTLOOK

Some examples of present experimental projects for a better WIMPs investigation are :

1) the Heidelberg-Moscow Collaboration, preparing a new detector with a ^{73}Ge enriched crystal [8].

2) a NaI(Tl) detector , of about 100 kg, being installed in the Gran Sasso Laboratory , by the BPRS Collaboration , mainly to start studying the annual modulation effect [10]. Moreover, a still better crystal purification is under way.

3) the CfPA group, preparing an "ice box" low activity cryostat to be installed in a shallow site at Stanford as a test bench for future R&D on bolometers. A second "ice box", to be installed deep underground, is also founded [19].

4) the Munich group, preparing a set of 4 x 250 g Sapphire (Al_2O_3) bolometers to be installed in Gran Sasso Laboratory by the end of 1996. They are focusing on phonon detection only and low mass WIMPs (< 20 GeV) [20].

5) the French Collaboration EDELWEISS, preparing more massive bolometers (≈ 1 kg Sapphire) with thin film sensors for athermal phonon detection, and 70 g Ge diodes with simultaneous detection of heat and ionization signals [21].

6) the Milano-Gran Sasso group, preparing a series of 20 (and then 100) TeO_2 bolometers , of 340g each, for both $\beta\beta$ and Dark Matter searches [23c].

5 . CONCLUSIONS

As discussed throughout this paper, at least up to now, **there is no "universal" detector** for WIMPs!

1) The **"reference"** detectors are the **classical Ge crystals** operating **at LN2 temperature.**

2) **Scintillators** have proved to be **well adapted** and **realistic** detectors for WIMPs search , mainly in view of an investigation of the **annual (and diurnal) modulation** effects. Further efforts to still reduce and/or reject residual background are in progress.

3) **Bolometers** represent the **most ambitious** developments, in particular by measuring **2 physical quantities** : Heat + Ionization or Heat + Scintillation. The remaining problems are a big challenge for the concerned groups.

4) **TPC** detectors, **superconducting grains** and **ancient mica** need further developments to reach a competitive level and challenge the other detectors already in operation.

REFERENCES

1. For a review of the Dark Matter problem and detection methods see, e.g. :
   - J.R. Primack, B. Sadoulet and D.Seckel, Ann. Rev. Nucl. Sci. 38 (1988) 751.
   - P.F. Smith and J.D. Lewin ,*"Dark Matter Detection"*, Phys. Rep. 187 (1990).
2. A. Bottino , et al., Astroparticle Physics **2,** 67 and 7 (1994).
3. M.W. Goodman and E.Witten , Phys.Rev. D 31(1985) 3059 .
4. A.K. Drukier et al, Nucl. Phys. B(Proc.Suppl.) 28A (1992) TAUP'91(Sept.)
5. D.O. Caldwell et al, Phys.Rev.Lett., 61(1988) 510.
6. D. Reusser et al., Phys. Lett. B 255 (1991) 143.
7 a) M.L.Sarsa (Zaragoza/USC/PNL) - TAUP93 Workshop at LNGS / INFN (Italy). 19-23 September 1993. Nucl. Phys. B (Proc. Suppl.) 35 (1994) 154.
   b) M.L.Sarsa (Zaragoza) *"Cold Dark Matter search with NaI scintillators at the Canfranc Underground Laboratory : looking for signal modulations"* Contribution at TAUP95 Conference (Toledo - September 1995).
   c) D.E.Di Gregorio (Buenos Aires/ Zaragoza/USC/PNL Collab.)*"Looking for diurnal modulation effects of WIMPs in the Southern Emisphere"* Contribution at TAUP95 Conference (Toledo - September 1995).
8. M. Beck (Heidelber-Moscow Collab.) *"The Heidelberg-Moscow experiment : Searching for Dark Matter"* Int. Workshop "The Dark Side of the Universe" at Tor Vergata Univ.(RomaII) June 1993;
   H.V. Klapdor *"Double Beta Decay and Dark Matter Results from the Moscow-Heidelberg Experiment"* contribution at TAUP95 Conference (Toledo - September 1995) and
   Y. Ramachers *"Phenomenological and Experimental activities for Dark Matter search in Heidelberg"* Int. Workshop "The Dark Side of the Universe" at Tor Vergata Univ.(RomaII) June 1995;
9. D.O. Caldwell et al. Phys. Rev. Lett. 65 (1990)1305.
10. P. Belli (BPRS Collab.) *"WIMPs search with low activity scintillators : status report"* Contribution at TAUP95 Conference (Toledo - September 1995), and C. Bacci et al.Astroparticle Phys. **2** (1994) 117 and E.Gaillard-Lecanu Doctor Thesis *"Détecteurs à cristaux scintillants pour la recherche de Matière Noire non-baryonique"*, June 1st, 1994.
    A. Incicchitti *"Dark Matter Search with low activity crystals by BPRS Collaboration"* Int. Workshop "The Dark Side of the Universe" at Tor Vergata Univ.(RomaII) June 1995.
11. G.J. Davies et al., Phys. Lett. B322 (1994)159 ; N.G.C. Spooner and Smith P.F., Phys. Lett. B314 (1993) 430; N.G.C. Spooner (Imp.College /Oxford/Rutherford Collab.) *"New Dark Matter Limits from the Boulby NaI detectors"*. Contribution at TAUP95 Conference (Toledo - September 1995);
    N. Spooner *"Dark Matter Limits from the NaI Detectors at Boulby"* Int. Workshop "The Dark Side of the Universe" at Tor Vergata Univ.(RomaII) June 1995 and Tariq Ali *"New Limits to Interaction Rates from the U.K. Dark Matter Experiment"* contribution to this Conference.
12. K. Fushimi et al., Phys. Rev. C 47, R425 (1993) and H. Ejiri et al., Phys. Lett. B317 (1993) 14.
13. A. Incicchitti (DAMA group) *"Particle Dark Matter Search with low activity scintillators : an update"* ROM2F, May 1995, contributed paper to the XXIV ICRC, Rome, August 1995 ; P. Belli et al. Nucl. Instr. and Methods, A336 (1993) 336, its contribution at TAUP95 Conference (Toledo - September 1995): see ref. 10, and its contribution to the Int. Workshop "The Dark Side of the Universe" at Tor Vergata Univ.(RomaII) June 1995 : *"Particle Dark Matter Search with Low activity Liquid Xenon Scintillators"*.
14. EDELWEISS Collab.(IPN Lyon, IN_2P_3, SACLAY, CdF, IAP, CSNSM) *"Calibration of a Ge crystal with nuclear recoils for the development of a Dark Matter detector"* Astroparticle physics 3 (1995) 361.
15. G. Gerbier et al., Phys. Rev. D 42 (1990) 3211.

16. K. Fushimi (Osaka/ Tokushima Collaboration) *"Search for Spin Coupled Dark Matter by means of Axial Vector Excitation of Nuclei "* Contribution at TAUP95 Conference (Toledo - September 1995) and *"Search for spin-coupled Dark Matter by means of Large Volume Scintillation Detectors" :* contribution to this Conference.
17. C. Bacci et al. Phys Lett. B 293 (1992) 460.
18. C. Bacci et al. Astroparticle Physics **2** (1994)13.
19. A. Lu (CfPA Collab.) talk given at at the XIXth "Rencontres de Moriond" at Villars-sur-Ollon (Switzerland), January 1994, and T. Shutt *"The Cryogenic Dark Matter Search experiment (CDMS)"* contribution to this Conference.
20. W. Seidel (MPI Munich / TUM Garching / Oxford Collaboration) *"Status of the CRESST Experiment"* Contribution at TAUP95 Conference (Toledo - September 1995), and L. Zerle *"The CRESS Dark Matter Search"* : contribution to this Conference.
21. D. Yvon (EDELWEISS Collaboration) *"First results of the EDELWEISS Experiment"* contribution at TAUP95 Conference (Toledo - September 1995) and V. Chazal (EDELWEISS Collaboration) *"Dark Matter Search with a Sapphire bolometer and Neutron background measurements in the Underground Laboratory of Modane"* contribution to this Conference.
22. M. Minowa et al., Nucl. Instr. and Methods in Physics Research, A327, (1993) 612.
23. a) M. Pavan (Milano Univ. and L.N.G.S.) *"Development of low temperature thermal detectors for Dark Matter : first measurements with different types of crystals"* . Int. Workshop "The Dark Side of the Universe" at Tor Vergata Univ. (Roma II) June 1993, and
    b) E. Fiorini (Milano Univ. and L.N.G.S.) *"Cryogenic Detection of Neutrinos, Dark Matter and Rare Processes"* Review talk at TAUP95 Conference (Toledo - September 1995).
    c) M. Pavan (Milano Univ. and L.N.G.S.) *"WIMPs search with TeO_2 thermal detectors : first results of the Milano-Gran Sasso experiment"* contribution to this conference.
24. A. Alessandrello et al., Nucl. Phys. B 28A (1992) 233, and ref.[18].
25. P. Belli et al. Nucl. Instr. & Methods in Physics Search A 357 (1995) 329.
26a) B. Chambon (IPN - Lyon) , private communication (on pure CsI).
    b) C. Bobin - doctor thesis *"Bolomètres massifs et détection de la matière noire non baryonique"* (IPN - Lyon) on a double bolometer CaF_2(Eu)-Sapphire. LYCEN T 9518 - 33-95 (January 1995).
27. O. Besida *"A time projection Chamber as a Dark Matter Detector"* Int. Workshop "The Dark Side of the Universe" at Tor Vergata Univ. June 1993; and Doctor Thesis : *"Etude d'un détecteur gaseux à ionisation pour la recherche de la Matière Noire de l'Univers."* March, 1994.
28. K. Buckland (San Diego) Berkeley workshop *"Strategies for the detection of Dark Matter Particles."* February, 1994.
29. K. Schmiemann *"The Orpheus Dark Matter Experiment"* - LTD6 workshop at Beatenberg/Interlaken - August 1995 and Int. Workshop "The Dark Side of the Universe" at Tor Vergata Univ.(RomaII) June 1995;
30. T. Girard (Univ. of Lisbon) *"The Lisbon-Zaragoza-Paris Dark Matter Search"* Contribution at TAUP95 Conference (Toledo - September 1995).
31. D.P. Snowden *"A unique signature for WIMPs in Ancient Mica"* contribution to this Conference.
32. J. Ellis and R.A. Flores, Phys Lett. B263 (1991) 259.
33. J. Engel, Phys Lett. B264 (1991) 114.

DÉTECTION DIRECTE DE PARTICULES CANDIDATES POUR LA "MATIÈRE SOMBRE"

Les expériences de détection de particules de matière sombre sont brièvement passées en revue. Les principaux problèmes rencontrés dans ce type de recherche sont soulignés et des stratégies envisageables pour les surmonter sont discutées.

DARK MATTER SEARCH WITH A 24 G SAPPHIRE BOLOMETER AND NEUTRON BACKGROUND MEASUREMENTS IN THE UNDERGROUND LABORATORY OF MODANE

Presented by Vérène Chazal
Institut de Physique Nucléaire de Lyon - 69100 Villeurbanne - France

EDELWEISS collaboration
and ISN Grenoble : R.Brissot, J.F.Cavaignac, A.Stutz

Abstract

The EDELWEISS collaboration proposes to look for the Dark Matter which can exist in WIMP form. In November 94, we obtained a preliminary result in the Underground Laboratory of Modane, with a 24 g sapphire bolometer. We measured an event rate of 25 evt/ kg/ keV/ day.

Another challenge of EDELWEISS is the elimination of the background neutron flux at the experimental site. We measured a fast neutron flux of 2.10^{-7} neut/ s/ cm^2. A thermic neutron flux study is outstanding, to make a global interpretation about the neutron environment in Modane.

1. Introduction

For 5 years, the EDELWEISS (Expérience pour Détecter les WImps En Site Souterrain) collaboration has been looking for - out of accelerator - the Dark Matter which can exist in WIMP (Weakly Interacting Massive Particle) form. These particles, by the supersymmetrie theory, have a mass of between 20 GeV and 1 TeV.

The two specificities of these particles are their low kinetic energy and their low interaction probability with ordinary matter. This characterizes the main part of our research, based on :

- the measure of very small signals, corresponding to very low energy transfer ($\approx$ keV) of the WIMP, by elastic scattering on the detector crystalline nucleus,
- the rare event rate $\approx 10^{-3}$ evt/ kg/ keV/ day, which imposes the elimination of all parasitic radiation from cosmic or radioactive origin.

2. Experimental installation

The L.S.M. (Laboratoire Souterrain de Modane) is situated in the Fréjus tunnel, at the French-Italian border. It is shielded by 1780 m of rock, 4400 m.w.e, which provides adequate protection against cosmic radiation.

The experiment has been installed in this site since September 1994. Isolation from microphonic noise and an antisismic plateform have been set up. All the material used have been specially selected : archeological lead of the Roman epoch, low activity lead and copper, and so on... An automated gas handling has also been developped, simplifing the cooling for the cooling and the maintenance of the cryostat. During measurements, there is a continuous flow of nitrogen, between the cryostat and the shield, to sweep out the radon gas.

3. Results

In November 94, we obtained a preliminary recoil spectrum with a 24 g sapphire bolometer (fig.1).

Running time : 160 hrs
T = 55-65 mK
Effective energy threshold ≈ 2 keV

with a NTD Germanium thermistor

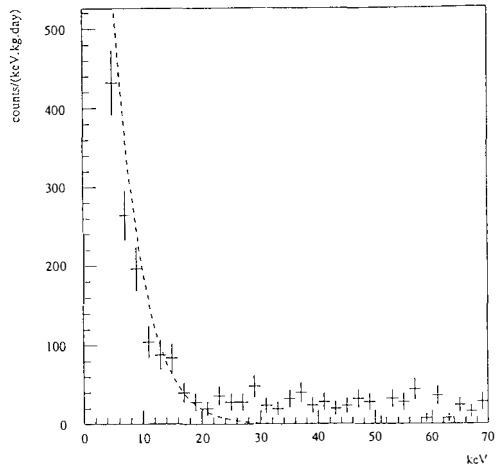


fig.1 : Experimental event rate (evt/ kg/ keV/ day)

Above 16 keV, we measured an event rate of 25 evt/ kg/ keV/ day. The dashed points are the calculated ten GeV WIMP spectrum.

From these results, we have been able to calculate an exclusion plot for axially and vector coupled WIMPs. Figure 2 represents the axially coupled WIMPs.

Al (94) : present experiment

Al (91) : 24 g sapphire bolometer
previous experiment [Cor 93]

Na (92) : NaI scintillator
experiment [Bac 92]

F : CaF_2 scintillator
experiment [Bac 92]

NaI (95) : NaI scintillator experiment,
with and without pulse shape
discrimination (PSD) [Ger 95]

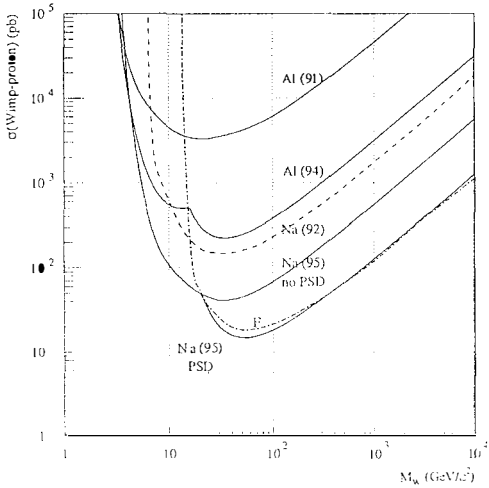


fig.2 : Exclusion plots for axially coupled WIMPs

This result (Al 94) is very encouraging, compared with the one obtained in 91 (Al 91) with the same bolometer in the L.S.M., but without special care concerning the cryostat. Other results have been added to the graph for comparaison.

But with 25 evt/ kg/ keV/ day, our results are two oders of magnitude higher than the most optimistic SUSY prediction.

4. Future program

Our next objects are :

- to upgrade the dilution refrigerator to have a base temperature of about 10 mK,
- to make measurements with simultaneous detection of heat and ionisation signals, with 70 g Ge diodes, for an active rejection of background radiation,
- to make new measurements with massive sapphire bolometers (200 g and 1 kg),
- to understand the origin of the residual radioactive background, with a 100 cm^3 at 77 K, inside the cryostat in the place of the bolometer,
- to continue the R & D program based on event discrimination.

5. Measurements of background neutron flux

Another challenge of EDELWEISS is the elimination of the background neutron flux. In direct searches for such dark matter, where one looks for the nuclear recoils produced by WIMPs scattering in the detector, neutrons will be the ultimate background noise source, since they produce the same signature as the hypothetical dark matter particles. So it is essential to minimize the energetic neutron flux at the experimental site, with a thorough study of the origin of the neutrons present.

A certain number of hypothesis have been considered : rock uranium-bearing ore fission, residual muons flux interaction especially in the lead, (α,n) reactions of natural radioactivity. The first object has been to measure with precision the flux and the energy distribution of these neutrons, and to compare this data to that estimated and coming from diverse origins, after analyse of the tunnel rock.

6. Experimental installation

The operation of the experimental device consisted of adapting the neutron detection module, developped for neutrinos detection by the BUGEY experiment [Ach 95]. The detector is composed of one parallelepiped (80 x 10 x 10 cm^3) joined to two photomultipliers. It is filled with an organic scintillator NE320, doped with 0.15 % ^6Li. A metallic shield of lead and copper protects the detector from residual gamma radiation. By elastic collisions on the hydrogen atoms of the scintillator , the neutron loses its energy (prompt pulse) and after a migration time ($\approx$ 20 μs), interacts

with a ^6Li nucleus to give a second pulse (delayed pulse). Five signatures allow identification of the neutron :

- the prompt pulse characteristic form (recoil proton), that we can discriminate from γ by their pulse shape (PSD). Thanks to good discrimination properties of the liquid, we can do a higher rejection of the γ induced background noise,
- the delayed pulse PSD, allowing the discrimination between the (α + t) charged particles of ^6Li + n reaction and γ,
- the energy of the (α + t) delayed pulses, with a magnitude of 4.8 MeV,
- the characteristic time ($t \approx 25$ µs) which separates the 2 signals corresponding to the neutron migration time. This is the typical thermalisation time of the neutron before it is captured,
- the thermal neutron migration distance.The difference between the prompt and delayed photomultipliers signals gives us the relative position of the interaction, with a resolution in position of about 20 cm.

7. Results

The whole installation was assembled in March 94 in the Fréjus tunnel, and the counting has taken place until December 94. The neutron rate has been estimated to 1.15 evt/ day in the detector, corresponding to an event rate of 2.10^{-7} neut / s / cm^2, for neutron energy greater than 1.5 MeV. This flux is comparable to a measured flux in the Gran Sasso by P.Belli et al [Bel 89], with a neutron energy greater than 2.5 MeV. The figure 3 represents the energy distribution, experimental (points) and simulated.

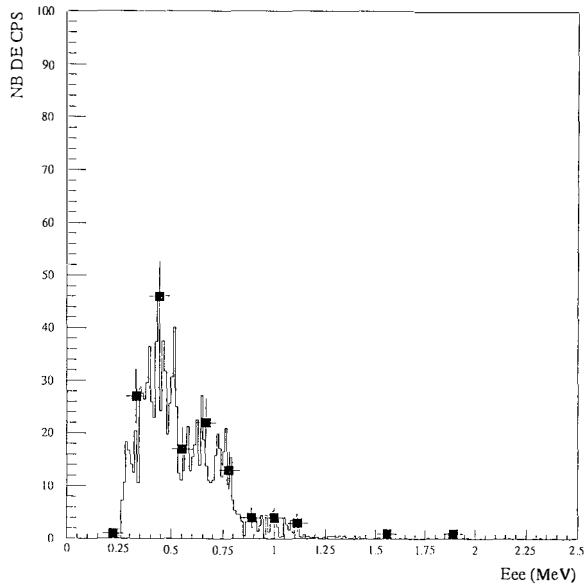


fig.3 : Energy spectrum

The experimental spectrum is obtained in electron equivalent energy, with a threshold effect on the first point. The simulation curve is obtained adding 3 weigted spectra of 2.5 MeV, 3 and 5 MeV neutrons energy, drawing on the lead surface. It seems that a same proportion of neutrons of 2.5 and 3 MeV constitutes a large part of the spectrum, 5 MeV neutrons allowing us to adjust the end of the spectrum. These are preliminary results.

In a second phase, from December 94 to May 95, we placed a paraffin layer, 30 cm thick all around the lead shield, to thermalise the fast neutrons. We obtained an event rate of 0.38 evt/ day in the detector.

In a third phase, from May 95 to December 95, we put an additional moderator, borax, between the copper and the detector. We found 0.27 evt/ day in the detector.

The interpretation of these results are outstanding, using simulation by the GEANT code. Particular attention is given to the evaluation of errors in these results.

A new experiment is going to be set up in the L.S.M., in December 95, with two small Helium-3 detectors, to measure the laboratory thermal neutron flux. In this way, we will be able to make a global interpretation about the neutron environment in Modane.

References

[Ach 95] : B.Achkar et al., Nucl. Phys. B 434 (1995) 503
[Bac 92] : C.Bacci et al., Phys. Lett. B 293 (1992) 460
[Bel 89] : P.Belli et al., Il Nuevo Cimento Vol.101 A, N.6 (1989) 959
[Cor 93] : N.Coron et al., Astron. Astrophys. 278 (1993) L31
[Ger 95] : G.Gerbier et al. , Proceedings of TAUP, Toledo (1995), to be published

Détection de Matière Noire avec un bolomètre Saphir 24 g et Mesure de l'Environnement Neutron au Laboratoire Souterrain de Modane

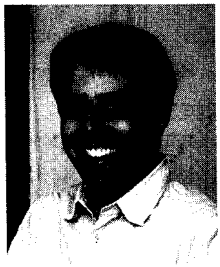
Résumé

La collaboration EDELWEISS se propose de rechercher la Matière Noire sous forme de WIMPs. En Septembre 1994, des premiers résultats ont été obtenu dans le Laboratoire Souterrain de Modane, avec un bolomètre saphir de 24 g. Nous avons mesuré un taux d'événements de 25 evt/kg/keV/j.

Un autre objectif d'EDELWEISS est l'élimination du bruit de fond neutron sur le site expérimental. Nous avons mesuré un flux de neutrons rapides de 2.10^{-7} neut/s/cm^2. Une étude du flux de neutrons thermiques est en cours, pour aboutir à une interprétation globale sur l'environnement neutron à Modane.

STATUS OF THE CRESST DARK MATTER SEARCH

L. Zerle[1]



for the CRESST collaboration

N.E. Booth[3], C. Bucci[4], M. Bühler[2], S. Cooper[1,3], F.v.Feilitzsch[2], P. Ferger[1], A. Gabutti[1], J. Höhne[2], E. Kellner[2], M. Loidl[1], O. Meier[1], U. Nagel[2] F. Pröbst[1], M.J.J.v.d.Putte[3], A. Rulofs[1], G.L. Salmon[3], W. Seidel[1] M. Sisti[1], L. Stodolsky[1], A. Stolovich[1]

[1] *Max-Planck-Institut für Physik, Föhringer Ring 6, D-80805 Munich, Germany*
[2] *Technische Universität München, Physik Department, D-85747 Munich, Germany*
[3] *Univeristy of Oxford, Physics Department, Oxford OX1 3RH, UK*
[4] *Laboratori Nazionali del Gran Sasso, I-67010 Assergi, Italy*

Abstract

We are preparing the CRESST experiment in LNGS to search for dark matter WIMPs using cryogenic detectors with superconducting phase transition thermometers. In the first stage of the experiment we plan to use four 250 g sapphire detectors with thresholds of 0.5 keV and resolutions of 0.2 keV at 1 keV. This will provide sensitivity to WIMP masses below 10 GeV, making our experiment complementary to other dark matter searches. In 1995 the construction of the main cryogenic components was completed and successfully tested. The installation of our equipment in LNGS has begun.

The CRESST (**C**ryogenic **R**are **E**vent **S**earch with **S**uperconducting **T**hermometers) experiment is designed to look for low-mass WIMPs as the dark matter of our Galaxy, using the cryogenic calorimeters developed in Munich. WIMPs could be detected via their elastic scattering on nuclei, giving nuclear recoil energies of about a keV for WIMP masses of a few GeV. Since low energy nuclei are very inefficient at producing ionization or scintillation, a low-threshold detector needs to act as a calorimeter, detecting the full nuclear energy. At the Max Planck Institute of Physics and the Technical University of Munich we have developed cryogenic calorimeters using superconducting phase transition thermometers. The thermometer is a small thin film of a superconducting material (tungsten) evaporated onto the surface of a sapphire crystal. The detector is run at a temperature ($\sim$ 15 mK) where the thermometer is in the middle of its transition between the normal and superconducting phases. Here its resistance is very sensitive to the small rise in temperature caused by a recoiling nucleus.

We have been able to detect 1.5 keV X-rays in a 32 g sapphire crystal with a resolution of 100 eV FWHM [1] in an above-ground laboratory. The local radioactivity already makes the background rate quite high in a detector of this size, so that the development of more massive detectors with this sensitivity requires a shielded underground site. However the model [2] which we have developed to describe the behaviour of our detectors leads us to expect that we can increase the detector mass by about an order of magnitude without losing much resolution. For the first phase of our dark matter experiment in Gran Sasso we are making 4 detectors of 250 g each, and hope to achieve a resolution of 200 eV at 1 keV and a threshold of 500 eV.

For a dark matter search experiment we need to combine the requirements of our detector (an operating temperature of $\sim$ 15 mK, provided by a dilution refrigerator) with the requirements of a low-background experiment (elimination of radioactivity). To eliminate cosmic ray background we perform the experiment in Gran Sasso Underground Laboratory (LNGS). Since a standard dilution refrigerator is made with various materials (stainless steel, indium vacuum seals) which are much too radioactive, we decided to separate the dilution refrigerator from the detector. The dilution refrigerator is based on a standard design from Oxford Instruments, with some modification to increase its mechanical strength. The detector will be placed in the "cold box", which hangs from the dilution refrigerator. This cold box is large enough to accommodate 100 kg of sapphire detectors.

The cold box is made of high-purity copper, with high-purity lead used for the vacuum seals. These materials are known to be among the best available for low radioactivity. As shown in the figure, the cold box will be surrounded by room-temperature shielding comprised of a 14 cm layer of high-purity copper and a 20 cm layer of lead. An internal shield serves to block any line of sight for radiation coming from the dilution refrigerator into the experimental volume.

A prototype cold box was designed and constructed in Munich, as was the gas handling system for the dilution refrigerator. The system was set up and its cryogenic properties tested at MPI in the fall of 1995. The temperature as measured in the cold box with ^{60}Co nuclear

orientation thermometry reached 6.8 mK, safely below that needed for the operation of our detectors. With this, a major milestone of the experiment has been reached.

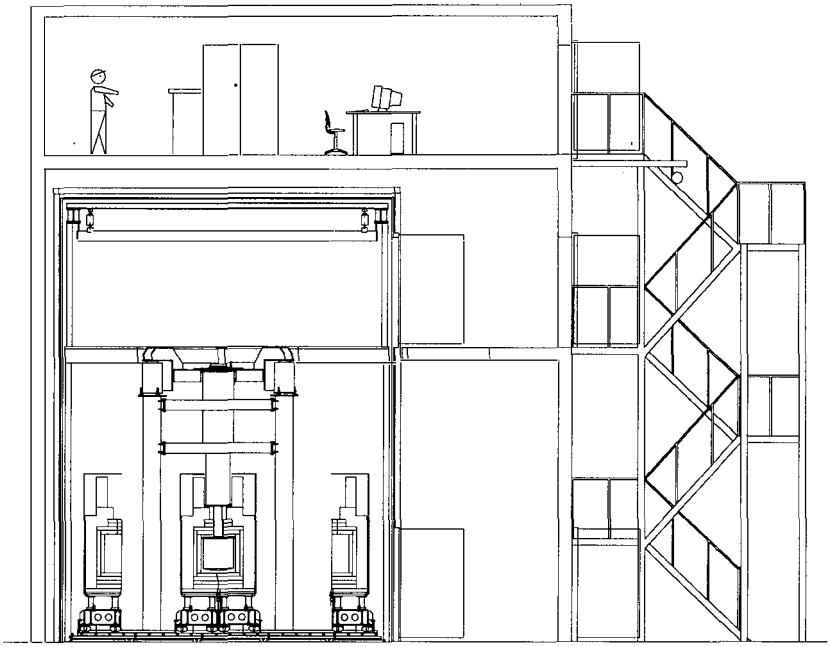


Figure 1: Cross section of CRESST building and experimental equipment in Hall B at Gran Sasso Laboratories. The dilution refrigerator (not drawn in this picture) is placed about 1 m above the "cold box" which is surrounded by low radioactivity shields

The copper for the shielding has undergone special handling [3] to assure its purity and to protect it from exposure to cosmic rays (which can transform copper into radioactive elements such as ^{60}Co). The copper was produced electrolytically by the Norddeutsche Affinerie in Dec. 1994 and rolled to the desired thickness a few weeks later. It was then placed in the cellar of a beer brewery near Munich, where it was shielded from cosmic rays by more than 10 m water equivalent. The final machining of the copper was performed by a company near Munich using special procedures for cleanliness. Each piece was only brought out of the brewery cellar for the few days needed for its machining, and then sealed in plastic, placed in a shipping crate, and returned to the cellar. The copper has now been transported to Gran Sasso and is stored in the underground laboratory. The total exposure to cosmic rays was 10 weeks.

It is not sufficient to use high-purity materials. Their surfaces must also be kept clean during use, and we have taken care to design our facilities in Gran Sasso to make this possible.

The Faraday cage which surrounds the experiment was chosen large enough so that all work on the low-background components of the experiment can be performed inside the cage. It is divided into two levels, with the upper level allowing access to the top of the cryostat and to the electronics. The lower level of the cage will be equipped as a clean room to protect the low-background components. The external lead and copper shields are in two closely fitting halves, each supported on a "wagon" so that the shielding can be opened without handling the individual pieces. In its retracted position (shown in the figure) the shielding is outside the dilution refrigerator support structure but still inside the clean room. Sufficient room is then available to disassemble the cold box which, since it consists of 5 shells, requires considerably more space as individual pieces. Entrance to the clean room will be through a changing room external to the Faraday cage.

We hope by the end of 1996 to get our first experience with a 250 g detector in our shielded environment in LNGS. Some adjustment of the detector parameters may then be needed to achieve optimum performance. These first measurements, which will be done using our prototype cold box, will also provide a first look at the radioactive backgrounds. The prototype was intended mainly as a test of the cryogenic techniques, and no attempt was made to limit the cosmic ray exposure or surface contamination. A new cold box will be fabricated using clean procedures similar to those used for the shielding. With this cold box, we hope in 1997 to achieve a good background rate and get our first significant data on dark matter.

References

[1] P. Colling, A. Nucciotti, C. Bucci, et al., "Low-energy X-ray detection in cryogenic detectors with tungsten thermometers", Nucl. Intr. Meth. A 354 (1995) 408.

[2] F. Pröbst, M. Frank, et al., "Model for Cryogenic Particle Detectors with Superconducting Phase Transition Thermometers", J. Low Temp. Phys. 100 (1995) 69.

[3] M. Biihler, L. Zerle, et al., "Status and Low Background Considerations for the CRESST Dark Matter Search", Nucl. Intr. Meth. A 370 (1996) 237.

WIMPS SEARCH WITH TeO_2 THERMAL DETECTORS: FIRST RESULTS OF THE MILANO-GRAN SASSO EXPERIMENT.

A. Alessandrello, C. Brofferio, D. V. Camin, P. Caspani, P. Colling+, O. Cremonesi, E. Fiorini, A. Giuliani, A. Nucciotti, M. Pavan, G. Pessina, E. Previtali, L. Zanotti
Dipartimento di Fisica dell'Università di Milano e Sezione di Milano dell'INFN, I-20133 Milano, Italy
+ *CEE fellowship in the Network on Cryogenic Detectors, under contract CHRX-CT93-0341.*

and

C. Bucci
Laboratori Nazionali del Gran Sasso, I-67010, Assergi (L'Aquila), Italy.

Abstract

Our group has performed - in the underground facility of Laboratori Nazionali del Gran Sasso - a Double Beta Decay Experiment using an array of four 340 g TeO_2 thermal detectors. The background spectrum (1548 hours) of the bolometer showing the best performances (i.e. 2 keV FWHM energy resolution and 13 keV energy threshold) has been analysed to search for WIMP signals. Lacking any evidence of WIMP interactions only limits on the spin independent cross section of Te and O nuclei could be obtained and are here reported.

1. Introduction

Although quite different in their physical purposes experiments aiming to detect WIMPs and experiments searching for neutrinoless Double Beta Decay (0ν-DBD) show to have very similar characteristics: in both cases large mass high resolution detectors are required, which have to be heavily shielded and placed underground. The main difference between the two experiments lays in the shape of the signal that in the case of 0ν-DBD search is a monochromatic line (usually in the MeV region) while in the case of WIMPs search it is a continuum spectrum (extending from few keV up to a hundred keV) whose only characteristic signature is the predicted seasonal modulation of the counting rate. As a consequence of these similar requirements DBD experiments may yield interesting results concerning WIMP flux and WIMP cross section.

2. Experimental details

Our group has performed a series of experiments searching for ^{130}Te DBD with low temperature TeO_2 thermal detectors[1]. Due to their high sensitivity to nuclear recoils, as well as to their high energy resolution, these devices are particularly suitable for WIMP detection, provided that a sufficiently low energy threshold is achieved. This last condition was satisfied for the first time during a DBD measurement performed with an array of four independent 340 g TeO_2 bolometers[2]. The array was operated in a dilution refrigerator installed in the underground facility of Laboratori Nazionali del Gran Sasso. The results here

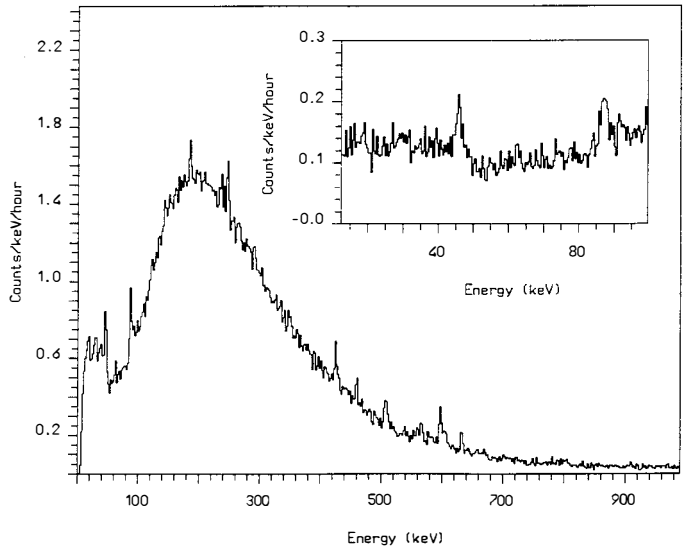


Figure 1: Background spectrum: in the insert the energy region between 13 keV and 100 keV.

reported have been obtained by analysing the 1548 hours background spectrum of one of the four bolometers of the array: the one having the best energy resolution and the lower energy threshold (Fig. 1). Below 1 MeV the spectrum of this detector is dominated by the bremsstrahlung of ^{210}Pb (a contaminant of the refrigerator external lead shield). The detector internal and external radioactive contaminations (mainly in U and Th) are responsible of several X and γ lines that are clearly evident in the background spectrum and that have been used for calibration. For dark matter analysis the region between the threshold, 13 keV, and 100 keV was considered. Here eight lines, the lowest at 27 keV the highest at 92 keV, allowed a careful control of energy resolution (2 keV FWHM) and linearity. The nuclear recoil quenching factor was determined in this same experiment by evaluating the detector response to ^{206}Pb recoils. This isotope is the daughter of ^{210}Po which decays by emitting a 5304 keV α particle and a 104 keV ^{206}Pb recoiling nucleus. The detector internal and external ^{210}Po contamination gives rise to two separate peaks in the spectrum. One is due to the internal contamination: both the α particle and the recoiling nucleus contribute to the signal so that this peak corresponds to the 5407 keV transition energy of ^{210}Po. The other peak is due to the external ^{210}Po contamination: the signal is produced by the α particle only and its energy is 5304 keV. The separation between the two peaks is 95.5 ± 1 keV, to be compared with the expected value of 103 keV. This difference, which is to be attributed to a different energy release between photons and nuclear recoils, is a measurement of the nuclear recoil quenching factor. In this experiment also a direct observation of ^{206}Pb recoils was made possible: by selecting only events in coincidence with an α signal (5304 keV peak) in one of the three nearby bolometers most background is rejected and the ^{206}Pb recoil peak becomes evident. Its position is 96 ± 3 keV in excellent agreement with the value obtained from the ^{210}Po peaks separation. By combining these two results a nuclear recoil quenching factor of 0.93 ± 0.03 is obtained.

The total detection efficiency, accounting for any probability of signal loss during data acquisition or data analysis, was estimated to be $(97 \pm 2)\%$.

3. Results

Lacking any signature of our events as a result of WIMP interactions we have conservatively assumed that all background counts were due to WIMP-nucleus elastic scattering processes. We have then compared the measured spectrum with the theoretical one foreseen for WIMP spin-independent interactions on Te or O nuclei, obtaining the exclusion plot shown in Fig. 2. The numerical parameters that have been used in data analysis are: a halo density of 0.3 GeV/cm^3, a galactic escape velocity of 2.67 10^{-3} c, and an Earth velocity in the galactic frame of 7.67 10^{-4} c. For the nuclear form factor the expression reported in ref. 3 has been assumed.

4. Conclusions

This result show the potentiality of thermal devices in detecting WIMPs. Actually the sensitivity of this experiment is limited not by detector performances but mainly by the high background counting rate. The next step in DBD search will be an experiment employing a 20 bolometer array (7 kg of TeO_2). Made with carefully selected low activity materials and surrounded by an ancient roman lead shield (without ^{210}Pb) this device should have a much reduced background and therefore a much increased sensitivity to WIMP interactions.

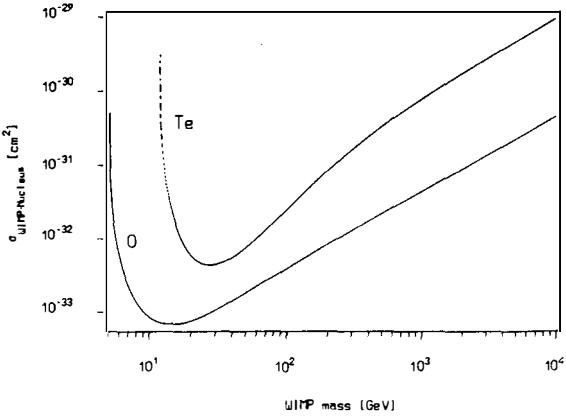


Figure 2: Exclusion plot for WIMPs cross section on Te and O.

References

1. A. Alessandrello et al., Phys Lett. B 335 (1994) 519
2. A. Alessandrello et al., NIM A 360 (1995) 363
3. J. Engel et al., Phys. Lett. B (1992) 119

Search for spin-coupled dark matter by means of large volume scintillators

H.Ejiri[a], K.Fushimi[b], K.Hayashi[a], R.Hazama[c], T.Kishimoto[c], N.Kudomi[c], K.Kume[c], T.Nitta[c], H.Ohsumi[c], and N.Suzuki[c]

a) Research Center for Nuclear Physics, Ibaraki, Osaka 567, Japan
b) Department of physics, The University of Tokushima, Tokushima, 770,Japan
c) Department of Physics, Osaka University, Toyonaka, Osaka 560, Japan

Abstract

Spin-coupled WIMPs particles were searched by for using large volume NaI and CaF_2 scintillators. Axial-vector (spin-coupled) excitation of ^{127}I by inelastic scattering of dark matter (DM) was studied to search for spin-coupled DM. A new stringent limit on the spin-coupled DM was obtained. A new detector system with CaF_2 was developed for studying elastic scattering of spin coupled WIMPs from ^{19}F.

The dark matter (DM) problem is a current important subject in astroparticle physics. Recent analysis of the cosmic microwave background by COBE, the large and small scale structure can be well fitted by mixed dark matter which consists of 60% of CDM, 30% of HDM and 10% of baryon [1]. WIMPs (Weakly Interacting Massive Particles) as a candidate of CDM has been extensively searched for by various detectors and methods. Direct measurement of WIMPs, which measures the interaction of WIMPs with nuclei, is hard because of the small flux, the weak interaction strength and the low energy depositl. So far, the direct measurements of WIMPs have been made by observing the nuclear recoil energy induced by the elastic scattering of WIMPs [2].

The elastic scattering of WIMPs was investigated by large NaI(Tl) scintillators [3] and CaF_2(Eu) scintillators [4]. Merits of these scintillators are large volume, low background, low energy threshold and the 100% isotope abundance of the finite-spin nuclei of ^{127}I, ^{23}Na and ^{19}F. The model dependence of the nuclear matrix element of spin-dependent elastic scattering is so large for ^{127}I and ^{23}Na nuclei [5]. The matrix element of ^{19}F can be calculated accurately [5]. In this paper, we first discuss the result by means of NaI(Tl) scintillator in ELEGANT V. Later, we discuss the new detector system ELEGANT VI which has been developed recently.

Inelastic nuclear excitation by spin-coupled WIMPs has been investigated by NaI(Tl) scintillator. The low-lying state with the excitation energy E^* is excited by the inelastic scattering of WIMPs. This excited state deexcites by the γ-transition, giving the energy deposit E^*. This energy is transferred to the electron energy in the detector and is converted fully to the light response $E_e = E^*$ without being reduced. The light output is the sum of the E^* for the γ-transition and the $f_I E_R$ for the recoil energy. Therefore the energy region to be looked for is shifted to the higher energy region (at the energy of excited state). There, the WIMPs signal can be fairly free from the low energy noise which lies in a few keV region.

The nuclear spin matrix element is difficult to calculate accurately because it is sensitive to the spin-isospin correlation. The spin matrix element for elastic scattering has recently been evaluated in terms of the odd-group model[5]. In the case of the spin-stretched transition to the excited state with spin $J' = J \pm 1$, the spin matrix element is related to the M1γ matrix element of $< A \mid M1 \mid A^* >$ as

$$< A^* \mid \mathbf{s} \mid A >= \sqrt{\frac{2J'+1}{2J+1}} \cdot \frac{1}{g_M} < A \mid M1 \mid A^* >, \qquad (1)$$

where g_M is the M1γ coupling constant given by $(e\hbar/2M)\ (3/4\pi)^{1/2}(g_s - g_\ell)/2$ and $< A \mid M1 \mid A^* >$ is obtained from the M1γ deexcitation (transition) rate.

The inelastic excitation of WIMPs from ^{127}I in the NaI(Tl) scintillator was studied by means of the huge NaI(Tl) scintillator of ELEGANT V. The detail of this system is given elsewhere [6]. Energy spectra, for the live time of 124days, from 17 modules of NaI detectors were summed for the final energy spectrum. Other 3 NaI modules were not used because of radioactive impurities. The observed energy spectrum indicates no distinct peak at 57.6keV.

The observed spectrum has been shown from the previous analysis [7] to be composed of the following background components: the β-ray spectrum from ^{210}Pb to the ground state of ^{210}Bi, the β-ray spectrum from ^{210}Pb to the first excited state of ^{210}Bi followed by the γ-ray emission, gaussian peaks of the X rays from U and Th chain isotopes and a continuous background. The energy spectrum was analyzed by means of the least square fit with these background components and possible component $S(E)$ due to the WIMPs inelastic excitation. The energy spectrum was folded with the energy resolution measured as $\Delta E = 7.8 + E/12$keV (FWHM).

A new stringent limit on the density of WIMPs was obtained. The upper limit of the WIMPs inelastic scattering is derived from the χ-square test of the fit as $S \leq 9.39 \times 10^{-2}$counts/day/kg with 90% CL. The solid line in Fig.1 shows the possible component of the WIMPs inelastic excitation. The original energy spectrum was shown in the previous report [8].

Using this upper limit on the event rate, one gets the limit on the local density ρ_0 of the WIMPs candidate, as shown in Fig.2. Since WIMPs needs to have enough energy to excite ^{127}I nuclei, the present method is quite sensitive to the heavy WIMPs candidates whose mass is heavier than several tens of GeV. It should be emphasized that the present limit is based on the observed nuclear spin matrix element derived from M1 gamma transition rate.

We have developed a new detector system, ELEGANT VI, which consists of CaF_2 scintillator-complexes. The detail of this detector system is shown in ref.[4]. The CaF_2(Eu) detector has great merits for detecting spin-coupled WIMPs: the low energy threshold ($\simeq$3keV), large mass fraction in CaF_2 crystal (49% , total mass of ^{19}F is 3.5kg) and the large nuclear matrix element for elastic scattering with small ambiguity. We study also double beta decay of ^{48}Ca by ELEGANT VI. The large volume ($45\times45\times200$mm^3) CaF_2(pure) crystals, which will be used in place of the CaF_2(Eu) crystals, contain 1.3mol of ^{48}Ca. The sensitivities of our detector for 0ν and $2\nu\beta\beta$ decays are $T^{0\nu}_{1/2} \simeq 1 \times 10^{23}$yr for one year measurement and $T^{2\nu}_{1/2} \simeq 2 \times 10^{20}$yr for 1000 hours measurement, respectively. The test experiment has been carried out in Osaka University (at sea level). Although the test was operated with a subset of the detector system, it has similar background rate as the BPRS Collaboration data in the Gran Sasso Underground laboratory [9]. The detector system will be installed at Oto cosmo observatory of 500mwe depth. The advantages of this laboratory are very low background levels of both cosmic ray and Rn. Cosmic rays with 4×10^{-3}m^{-2}sec^{-1}, and neutrons with 4×10^{-1}m^{-2}sec^{-1} are compara-

ble to those at Kamioka underground laboratory. The Rn content with $10 Bq \cdot m^{-3}$ is $2 \sim 3$ orders of magnitudes smaller than that at Kamioka[10].

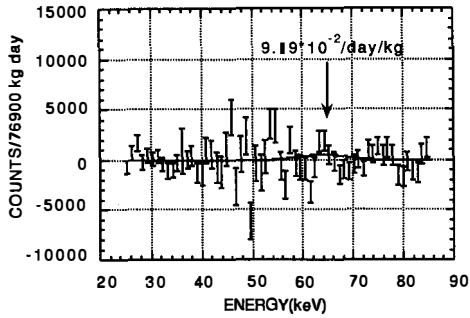


Figure 1: The subtracted energy spectrum in the low energy region obtained by the 17 NaI(Tl) scintillator. The gaussian peak shows possible spectrum of the inelastic scattering with the given event rate, corresponding to the 1.29σ(90% CL) limit.

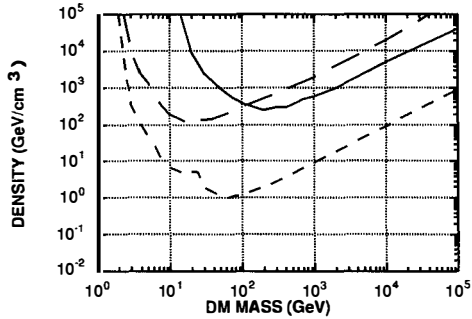


Figure 2: The exclusion plot with the 90% CL for the density of the Dirac neutrino (lower dashed) and the Majorana neutrino (upper dotted). The solid line indicates the present limit. The regions above the lines are excluded from the main component of WIMPs. The dash-dotted line indicates the critical density in the vicinity of the earth ($0.3 GeV cm^{-3}$).

The authors thank Professors K.Okada and E.Takasugi for valuable discussions. They owe much to the ELEGANT group at Osaka University for their help in operating ELEGANT V, and the Mitsui Metal Industry Ltd. and ICRR for the Kamioka underground laboratory. This work was supported by the Grant-in-Aid of Scientific Research, Ministry of Education, Science and Culture, Japan.

References

[1] G.F.Smoot et al., ApJ **396** (1992) L1; E.L.Wright et al., ApJ **396** (1992) L13.

[2] J.R.Primack et al., Ann.Rev.Nucl.Sci., **38** (1988) 751.
P.F.Smith et al., Physics Reports, **187** (1990) 203;
and papers appears in these proceedings.

[3] K.Fushimi, H.Ejiri, H.Kinoshita, N.Kudomi, K.Kume, K.Nagata, H.Ohsumi, K.Okada, H.Sano and J.Tanaka, Phys.Rev., **C47** (1993) R425.

[4] R.Hazama, H.Ejiri, K.Fushimi, T.Kishimoto, N.Kudomi, K.Kume, T.Nitta, H.Ohsumi and N.Suzuki, Proc. of the IV International Symposium on Weak and Electromagnetic Interactions in Nuclei (WEIN 95), Osaka, Japan, June 1995, World Scientific, p635.

[5] J.Ellis and R.Flores, Nucl.Phys. **B307** (1988) 883; Phys.Lett. **B263** (1991) 259.

[6] H.Ejiri et al., Nucl.Instrum.Meth. **A302** (1991) 304; Phys.Lett. **B258** (1991) 17.

[7] H.Ejiri et al., Phys.Lett. **B282** (1992) 281; Phys.Rev. **C44** (1991) 502.

[8] H.Ejiri, K.Fushimi and H.Ohsumi, Phys.Lett. **B317** (1993) 14.

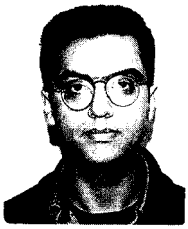
[9] C.Bacci et al. BPRS Collaboration, Astrop.Phys. **2** (1994) 117.

[10] H.Ejiri, Proc. of the IV International Symposium on Weak and Electromagnetic Interactions in Nuclei (WEIN 95), Osaka, Japan, June 1995, World Scientific, p687.

NEW LIMITS TO INTERACTION RATES FROM THE UK DARK MATTER EXPERIMENT

T.Ali*, J.J. Quenby*, T.J. Sumner*, A. Bewick*, J.P. Li*, D. Shaul*, N.J.T. Smith*, W.G. Jones*, G.J. Davies*, C.H. Lally*, M.K. Joshi*, P.F. Smith[1], G.J. Homer[1], G.T.J. Arnison[1], J.D.Lewin[1], G.J. Alner[1], M.J.T. van den Putte[1], N.J.C. Spooner[2], J.C. Barton[3], P.R. Blake[4].

* Blackett Laboratory, Imperial College of Science Technology & Medicine, London. SW7 2BZ. UK.
[1]Particle Physics dept., Rutherford Appleton Laboratory, Chilton, Didcot, Oxon. OX11 0QX. UK
[2]Physics Dept., University of Sheffield, Sheffield. S3 7RH. UK
[3]Physics Dept., Birckbeck College, London. WC1E 7HX. UK
[4]Physics Dept., University of Nottingham, Nottingham. NG7 2RD. UK



ABSTRACT

Direct detection of Weakly Interacting Massive Particles (WIMP's) is possible if they have a scattering cross-section with nucleons comparable to that of a neutrino. Low background underground detector searches are underway worldwide and we report here on an improvement of an order of magnitude in our published limits to the spin-dependent dark matter interaction rate. A 2σ upper limit to the rate of 0.14 events/kg.keV.day (at 7-10 keV equivalent photon energy) is achieved by use of event time profile information to reject most of the radioactive background.

Recent results from the microlensing community appear to show that <30% of the Galactic halo can be baryonic[1]. Thus the motivation for a non-baryonic dark matter search remains as strong as ever and is discussed elsewhere in these proceedings. Many groups worldwide are engaged on this quest and we presented our first results from a room temperature NaI scintillator experiment at a previous Moriond conference[2]. We report here an order of magnitude improvement over our latest published result[5] in the upper limit to the interaction rate achieved by using event profile information to reject background noise.

Experimental Details

The experiment is located at a depth of 1100m (3000 mwe) underground in a working mine near Boulby, England. Passive shielding is effected by means of lead and copper castles or by lowering the detector inside a pure copper vessel into a 6m diameter, 6m deep tank of recirculating pure water. All shielding materials have been stored underground for many years.
The results discussed are from data acquired over a 6 month run with a 6.2 kg NaI(Tl) crystal located in the water tank. The scintillation event readout using 2 low-noise EMI 9625A photomultipliers viewing opposite crystal faces through 30 cm silica light guides. Each coincident event is digitised and recorded from both PMTs. The photoelectron yield in this configuration was measured to be ~1.7/keV. A real time environmental monitoring system allows us to record the detector temperature.
The resulting low energy background spectrum is shown in fig1. Note that a rate of ~1.7 counts/keV/kg/day is measured for the 7-10 keV equivalent photon energy interval. Since this energy range represents a minimum in the background versus energy loss curve of fig1 and is near the detector threshold, we can obtain a best dark matter limit from this point on the curve.

Calibration

We expect a WIMP event signal to be similiar in its time decay characteristics to that produced by a neutron scattering in the crystal. Experimentally we observe two distinct event decay time constant distributions for neutrons and gamma rays with the separation diverging towards higher energy. As the residual neutron activity is negligable in the proximity of the experiment, it is possible to detect statistically the presence of a small fraction of dark matter

recoil events in the presence of a beta and gamma background by using an event decay time discrimination scheme.
Calibration runs in the water tank with a ^{252}Cf source emitting gammas and neutrons and a ^{57}Co source emitting gammas only yielded the two time constant distributions for the 7-10 keV energy range shown in fig2. A separation of the mean decay time constant between the two species of ~26ns can be seen. All time constants have been determined by a least-squares-fit of the event profile to a single exponential.

Results and Limits

We can compare a time constant distribution derived from the measured spectrum of fig1 with the calibration curves shown in fig2 for the 7-10 keV energy range. Corrections have been applied to the background spectrum for Poisson statistics of photoelectron generation and the probability of coincidentally triggering both PMT channels. In addition a correction for a secular reduction in light collection efficiency amounting to 2.5%/month and an empirical time constant temperature coefficient of 1.3%/degree K for a ±1K temperature drift have been made.
In fig2 data from the observed background spectrum is shown plotted and a near coincidence with the gamma calibration spectrum is seen. To 90% C.L. the difference between the means of these time constant distributions is 2ns, thus only 2/26 (0.008) of the observed 1.7 events/keV/kg/day count rate could be due to dark matter particles (see[3)4)]). The resultant 0.14 events/keV/kg/day is taken to represent the upper limit to dark matter interactions arising from this experiment.

We can relate the observations to a differential event rate/keV/kg/day by[5)];

$$dR/dE_R = (R_o/E_or)(\pi^{0.5}/4y)\ [erf(x+y)-erf(x-y)]\ [F(E_R)]^2 \qquad (1)$$

where F is the form factor, R_o is the reduced total event rate in counts/kg/day, $E_o=1/2M_d(v_o)^2$, M_d and M_t is the mass of the dark matter particle and the target nucleus respectively, $x=(E_R/E_o)$, $y=(v_E/v_o)$, $r=4(M_dM_t)/(M_d+M_t)^2$.
By using the 90% C.L. experimental limit in equation 1 we can solve for the normalised rate R_o/r. The resulting limit as a function of the particle mass is shown in fig3 for the spin dependent interaction. Also shown are our previous results for a 1.3kg detector[5)] and the Germanium (~8% ^{73}Ge isotope) limit[6-9)].

The limit obtained with our pulse shape discrimination scheme clearly represents an improvement on previous results.
We can also relate R_0/r to a total interaction cross section and an equivalent cross section for a single proton by using the appropriate spin factors for Na and H[3]. This yields a value of ~3pb in the 7-10 keV range.

Conclusion

We have demonstrated an order of magnitude improvement on our previous published spin-dependent limits by noise rejection using event decay time characteristics. A detailed statistical analysis of all the energy bins in the background spectrum yielding a small improvement on the above limit is in preparation[4].
We envisage further gains in sensitivity will be achieved by scaling up the detector size and with longer running times. We have initiated a development programme to produce large area photon-counting avalanche photodiodes. Deployment of these will result in an increased light collection efficiency, lower energy threshold and with a segmented array geometry, provide a further discrimination channel. In addition alternative routes to our goal including cooled NaI and liquid xenon are under investigation.

Acknowledgements
We thank Cleveland Potash LTD for the Boulby facilities, the support teams at ICSTM and RAL. STARLINK computing facilities have been used at ICSTM and PPARC provide funding for the experiment.

References
1) E.I. Gates, G. Gyuk and M.S. Turner. Phys. Rev. Lett. 74 (1995).
2) T.J. Sumner et al. Particle physics, atomic physics and gravitation (proc.14th Moriond workshop) eds Tran Thanh Van et al, Editions Frontiérs (1995).
3) J.J. Quenby et al. Astroparticle Physics, submitted (1996).
4) P.F. Smith et al. Phys. Lett. B in prep (1996).
5) J.J. Quenby et al Phys. Lett. B351 (1995).
6) D.O. Caldwell et al. Phys Rev. Lett. 61 (1988).
7) D. Reusser et al. Phys. Lett. B255 (1991).
8) A. Drukier et al. Nucl. Phys. B.(Proc. Suppl.) 28A (1994).
9) E. Garcia et al. Phys. Rev. D51 (1995).

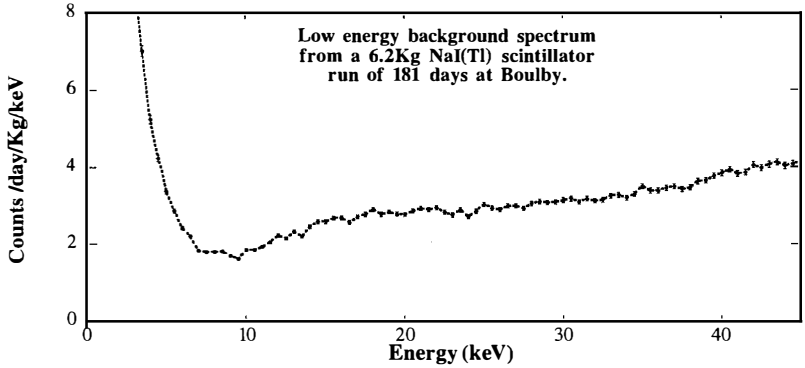


Figure 1: Low energy pulse height spectrum from a 6.2kg NaI(Tl) crystal

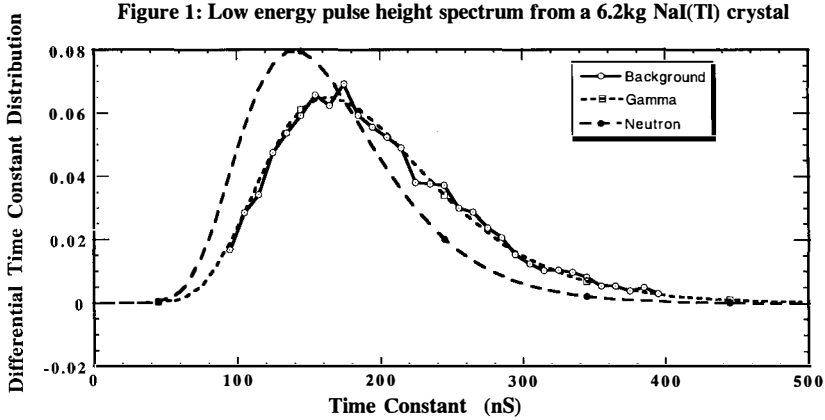


Figure 2: Normalised differential time-constant distributions for the 7-10 keV energy loss interval

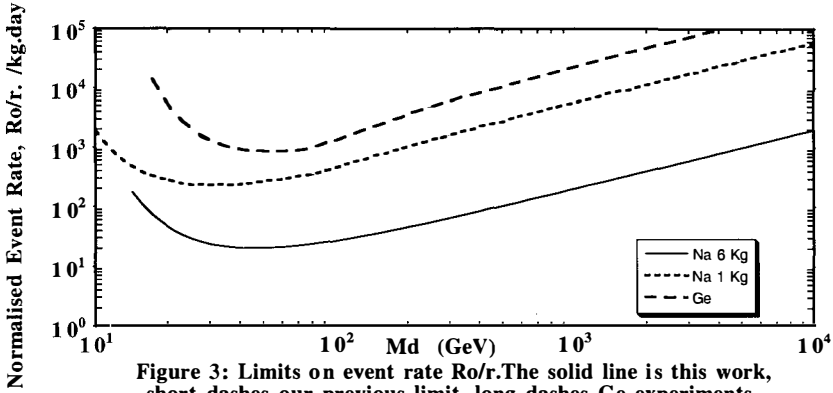


Figure 3: Limits on event rate Ro/r.The solid line is this work, short dashes our previous limit, long dashes Ge experiments

Indirect Detection of Supersymmetric Dark Matter

L. Bergström

Department of Physics
Stockholm University
Box 6730, S-113 85 Stockholm
Sweden
and
Department of Theoretical Physics
Uppsala University
Box 803
S-751 08 Uppsala, Sweden

Abstract

Indirect methods proposed for detecting supersymmetric dark matter are discussed. If the dark matter halo of our galaxy consists of the lightest supersymmetric particles, assumed to be neutralinos, their annihilation in the galactic halo may produce observable fluxes of positrons, antiprotons and gamma rays. In addition, neutralinos may be captured and enriched in the centre of the Sun or the Earth. The most promising methods seem to be the detection of a gamma ray line from the halo of energy equal to the neutralino, and detection of high energy neutrinos from the centre of the Sun or the Earth, where new neutrinos telescopes may either detect or put useful limits on these dark matter candidates.

1 Introduction

The mismatch between dynamical estimates of the mass density of the universe, especially on large scales, and the upper bound on baryonic density provided by nucleosynthesis constraints seems to point to the existence of non-baryonic dark matter. The lightest supersymmetric particle, the neutralino, may if it exists have just the right properties to explain the dark matter problem.

Why should then nature be supersymmetric? There are several answers to this question, none unfortunately very compelling, but together they may give a flavour of why particle theorists have become so attracted to the idea of supersymmetry.

First of all, the algebra between two supersymmetry generators

$$[\delta(\epsilon_1), \delta(\epsilon_2)] = \frac{i}{2}\bar{\epsilon}_2\gamma^\mu\epsilon_1 P_\mu \tag{1}$$

has the remarkable feature of linking the supersymmetry generators δ of the fields of the theory to the generator P^μ of the Poincare group. Besides being elegant, this may allow gravity to be unified with the other interactions (perhaps via superstrings).

The second strong reason for believing in supersymmetry has to do with the unnaturalness of the usual (non-supersymmetric) standard model in the sense that the weak interaction scale is some seventeen orders of magnitude lower than the Planck scale (the scale where gravity has to be treated quantum-mechanically), despite the fact that no symmetry is present to protect that low mass scale. This presents a tremendous fine-tuning problem, since radiative corrections would immediately drive the Higgs mass, and therefore the electroweak scale, towards the Planck scale unless fine-tuned counterterms are introduced order by order in perturbation theory. Supersymmetry, on the other hand, is "nice" to hierarchies: once there in lowest order, the susy corrections come with alternating signs between the fermions and the bosons in the theory and therefore do not spoil the hierarchy. The postulation of supersymmetry at a high mass scale with soft breaking terms (at the Grand Unification or even the Planck scale) may through radiative effects and renormalization group flow of the parameters generate an exponential hierarchy that explains the disparate lowest order mass scales of the effective theories.

The third reason is the fact that although Grand Unification is such an attractive possibility, the proton does not want to decay, and the gauge coupling constants seem to unify better at high energies if intermediate mass scales (such as new thresholds given by susy particles) are present [3].

Perhaps the strongest reason for susy is a very indirect one, and one that has only recently been appreciated: the remarkably beautiful new duality properties relating strong and weak coupling of supersymmetric theories, discovered by Seiberg and Witten [1]. This work has led to a phenomenal succession of interesting developments also in string theories and their higher-dimensional counterparts, once again reviving the hope of finding a unique "theory of everything".

With all these particle physics motivations for supersymmetry, the fact that supersymmetric models also can produce natural dark matter candidates comes as a nice extra bonus. (At this point, one may even be tempted to appeal to the populistic argument for the existence of supersymmetry: 1500 theorists and 5000 experimentalists cannot be wrong!)

Thus, the particle dark matter candidate we will focus on is the lightest supersymmetric particle χ, assumed to be a neutralino, i.e. a mixture of the supersymmetric partners of the photon, the Z^0 and the two neutral CP-even Higgs bosons present in the minimal extension of the supersymmetric standard model (see, e.g. [2]). The attractiveness of this candidate stems

from the fact that its generic couplings and mass range naturally gives a relic density close to the critical one. (For a recent review of supersymmetric dark matter, see Ref. 4.)

2 The minimal supersymmetric extension of the standard model

The minimal $N = 1$ supersymmetric extension of the standard model is defined by, besides the particle content and gauge couplings required by supersymmetry, the superpotential (for details, see Ref. 5)

$$W = \epsilon_{ij}\left(-\hat{\mathbf{e}}_R^* \mathbf{Y}_E \hat{\mathbf{l}}_L^i \hat{H}_1^j - \hat{\mathbf{d}}_R^* \mathbf{Y}_D \hat{\mathbf{q}}_L^i \hat{H}_1^j + \hat{\mathbf{u}}_R^* \mathbf{Y}_U \hat{\mathbf{q}}_L^i \hat{H}_2^j - \mu \hat{H}_1^i \hat{H}_2^j\right). \tag{2}$$

Breaking of supersymmetry has to be present (since no supersymmetric particles have as yet been detected, and supersymmetry requires particles and sparticles to have the same mass). This can be achieved by a soft supersymmetry-breaking potential which does not re-introduce big radiative mass-shifts.

Often a simplifying unification assumption is used for the gaugino mass parameters,

$$\begin{aligned} M_1 &= \tfrac{5}{3}\tan^2\theta_w M_2 \simeq 0.5 M_2, \\ M_2 &= \tfrac{\alpha_{\rm EW}}{\sin^2\theta_w \alpha_s} M_3 \simeq 0.3 M_3. \end{aligned} \tag{3}$$

Electroweak symmetry breaking is caused by both neutral Higgs fields in the the theory, H_1^1 and H_2^2, acquiring vacuum expectation values,

$$\langle H_1^1 \rangle = v_1, \qquad \langle H_2^2 \rangle = v_2, \tag{4}$$

with $g^2(v_1^2 + v_2^2) = 2m_W^2$, with the further assumption that vacuum expectation values of all other scalar fields (in particular, squark and sleptons) vanish. This avoids color and/or charge breaking vacua.

The minimization conditions of the Higgs potential allow one to trade two of the three arbitrary parameters entering the Higgs potential μ_1^2, μ_2^2 and $B\mu$ for the Z boson mass $m_Z^2 = \frac{1}{2}(g^2 + g'^2)(v_1^2 + v_2^2)$ and the ratio of vevs $\tan\beta = v_2/v_1$. The third parameter can further be reexpressed in terms of the mass of one of the physical Higgs bosons.

Choosing as independent parameter the mass m_A of the CP-odd Higgs boson, the masses of the other Higgs bosons are given by

$$\mathcal{M}_H^2 = \begin{pmatrix} m_A^2 \cos^2\beta + m_Z^2 \sin^2\beta + \Delta\mathcal{M}_{11}^2 & -\sin\beta\cos\beta(m_A^2 + m_Z^2) + \Delta\mathcal{M}_{12}^2 \\ -\sin\beta\cos\beta(m_A^2 + m_Z^2) + \Delta\mathcal{M}_{21}^2 & m_A^2 \sin^2\beta + m_Z^2\cos^2\beta + \Delta\mathcal{M}_{22}^2 \end{pmatrix} \tag{5}$$

$$m_{H^\pm}^2 = m_A^2 + m_W^2 + \Delta_\pm. \tag{6}$$

The quantities $\Delta\mathcal{M}_{ij}^2$ and $\Delta_\pm$ are the one-loop radiative corrections coming from virtual (s)top and (s)bottom loops. Diagonalization of $\mathcal{M}_H^2$ gives the two CP-even Higgs boson masses, $m_{H_{1,2}}$, and their mixing angle.

The neutralinos $\tilde{\chi}_i^0$ are linear combination of the neutral gauge bosons $\tilde{B}$, $\tilde{W}_3$ and of the neutral higgsinos $\tilde{H}_1^0$, $\tilde{H}_2^0$. In this basis, their mass matrix

$$\mathcal{M}_{\tilde{\chi}^0_{1,2,3,4}} = \begin{pmatrix} M_1 & 0 & -\frac{g'v_1}{\sqrt{2}} & +\frac{g'v_2}{\sqrt{2}} \\ 0 & M_2 & +\frac{gv_1}{\sqrt{2}} & -\frac{gv_2}{\sqrt{2}} \\ -\frac{g'v_1}{\sqrt{2}} & +\frac{gv_1}{\sqrt{2}} & 0 & -\mu \\ +\frac{g'v_2}{\sqrt{2}} & -\frac{gv_2}{\sqrt{2}} & -\mu & 0 \end{pmatrix} \tag{7}$$

can be diagonalized analytically to give four neutral Majorana states,

$$\tilde{\chi}_i^0 = Z_{i1}\tilde{B} + Z_{i2}\tilde{W}^3 + Z_{i3}\tilde{H}_1^0 + Z_{i4}\tilde{H}_2^0, \quad (8)$$

the lightest of which, to be called χ, is then the candidate for the particle making up (at least some of) the dark matter in the universe.

When using the minimal supersymmetric standard model in calculations of relic dark matter density, one should of course make sure that all accelerator constraints on supersymmetric particles and couplings are imposed. In particular, the measurement of the $b \to s\gamma$ process at the Cornell accelerator [7] is providing important bounds. Also, the recent 130 GeV run at LEP has given significant new constrints (e.g. raising the lower bound of the chargino to around 65 GeV).

3 Indirect detection methods

Besides the possibilities of direct detection of supersymmetric dark matter (see the talk by P. Gondolo at this Workshop), one should also consider indirect detection methods of neutralino annihilation in the galactic halo.

In principle, all stable particles produced in annihilation processes in the halo can serve as signatures of the dark matter candidate. However, electrons and protons are much too abundant in the ordinary cosmic rays to be useful. Much lower background fluxes, and therefore greater potential for detection of an additional component, are present in the case of positrons and antiprotons. The problem is still that the signal does not have a distinct signature. This is due to the fact that the primary annihilation processes are into quarks, heavy leptons, gauge bosons and Higgs particles, whereas the positrons and antiprotons are secondary or tertiary decay products. The reason for the small direct coupling to electron-positron pairs is the Majorana nature of the neutralino coupled with the fact that the annihilation takes place essentially at rest. Selection rules then force the coupling to fermions to be proportional to the fermion mass.

There has recently been a new balloon-borne detection experiment [8], with increased sensitivity to eventual positrons from neutralino annihilation, but also this is far from being able to put stringent limits on such dark matter annihilation processes. Antiprotons could for some supersymmetric parameters constitute a useful signal [9], but probably space experiments [10] would be needed to disentangle a low-energy signal from the smooth cosmic-ray induced background. A problem that plagues estimates of the signal strength of both positrons and

antiprotons is the uncertainty of the galactic propagation model, and solar wind modulations.

4 Gamma rays

With these problems of positrons and antiprotons, one would expect that problems of gamma rays and neutrinos are similar, if they only arise from secondary decays in the annihilation process. This is certainly true if that were the only source of gammas and neutrinos. However, neutrinos can escape from the centre of the Sun or Earth, where neutralinos may have been gravitationally trapped and therefore their density

enhanced [11]. Gamma rays may result from loop-induced annihilations $\chi\chi \to \gamma\gamma$ [12] or $\chi\chi \to \gamma Z^0$ [13].

Theses processes may still be difficult to estimate because of uncertainties in the supersymmetric parameters, cross sections and halo density profile. However, in contrast to the other proposed detection methods they have the virtue of giving a very distinct, "smoking gun" signal: high-energy neutrinos from the centre of the Earth or Sun, and monoenergetic photons

with $E_\gamma = m_\chi$ from the halo (the signal in that case is expected to be strongest in the direction of the galactic centre). In addition, the detection of a gamma line signal could enable a map of the density profile of the dark matter halo.

To illustrate the latter point, let us consider the characteristic angular dependence of the gamma-ray intensity from neutralino annihilation in the galactic halo [14]. Annihilation of neutralinos in an isothermal halo with core radius a leads to a gamma-ray flux of

$$\frac{d\mathcal{F}}{d\Omega} = \frac{\sigma_{\chi\chi\to\gamma\gamma} v}{4\pi m_\chi^2} \int_0^\infty \rho^2(r)\, dr(\vartheta) \simeq (2\times 10^{-8}\mathrm{m}^{-2}\mathrm{s}^{-1}\mathrm{sr}^{-1}) \frac{(\sigma_{\gamma\gamma} v/\, 10^{-30}\,\mathrm{cm}^3\,\mathrm{s}^{-1})(\rho_\chi^{0.4})^2}{(m_\chi/\,10\,\mathrm{GeV})^2}\, I(\vartheta) \quad (9)$$

where ϑ is the angle between the line of sight and the Galactic center, $r(\vartheta)$ is the distance along that line of sight, and $\rho_\chi^{0.4}$ is the local neutralino halo density in units of 0.4 GeV cm^{-3}. The quantity $\sigma_{\gamma\gamma} v$ is the cross section times relative velocity v for annihilation of neutralinos into two photons. We remind of the fact that since the neutralino velocities in the halo are of the order of 10^{-3} of the velocity of light, the annihilation can be considered to be at rest. The resulting gamma ray spectrum is a line at $E_\gamma = m_\chi$ of relative linewidth 10^{-3} which in favourable cases could have a chance to stand out against background.

The calculation of the $\chi\chi \to \gamma\gamma$ cross section is technically quite involved with a large number of loop diagrams contributing. In fact, a full calculation still remains to be done (the most complete one up to date is that of [15]). However, it is quite easy to get an order of magnitude estimate of the various contributions. Loops with fermions and sfermions in the intermediate state generally contribute an amplitude of order $\alpha^2 m_\chi / m_{\tilde{f}}^2$. Since the different contributions all have to be added coherently, there may be cancellations or enhancements, depending on the supersymmetric parameters.

A potentially important contribution, especially for neutralinos that contain a fair fraction of Higgsino components, is from W^+W^- intermediate states. This is also the dominant loop for the even more important $Z^0\gamma$ final state for very massive neutralinos [13]. Writing

$$\mathcal{A}_{\gamma\gamma} = \frac{Ce^2}{4\pi}\epsilon[\mu\nu q_1 q_2]\,(M_{re} + iM_{im})\,, \quad (10)$$

one finds [13]

$$M_{im} = \beta^2 \log\left(\frac{1+\beta}{1-\beta}\right), \quad (11)$$

where $\beta = \sqrt{1-(m_W/m_\chi)^2}$.

In the limit of large χ masses (above, say, 500 GeV) this becomes

$$M_{im}^{LL} = \log(\frac{s}{m_W^2}), \quad (12)$$

$$M_{re}^{LL} = \frac{1}{2\pi}\log^2(\frac{s}{m_W^2}), \quad (13)$$

where $s = 4m_\chi^2$. Non-leading corrections to this result have recently also been calculated [15].

In Fig. 1, we show the gamma ray line flux given in a scan [5] of supersymmetric models consistent will all experimental bounds (including $b \to s\gamma$). The open points are from a "standard" scan in the parameter space letting μ, M_2, tanβ and m_{H_3} vary at random between generous bounds. The filled points are from a "special" scan, where as an extra condition a spin independent cross section on nucleons greater than 5% of the Dirac cross section has been required. As can be seen, this condition has little effect on the gamma rates (the two populations

of points cannot be easily distinguished), in contrast to the case of neutrinos discussed below (see Fig. 2) and to direct detection rates [5]).

It can be seen that some models give rates that could be probed by a next generation of air Cherenkov telescopes [16], which would have an area of up to several tens of thousand square metres.

5 Neutrinos

Another promising indirect detection method is to use neutrinos from annihilations of neutralinos accumulated in the centre of the Sun or Earth. This will be a field of extensive experimental investigations in view of the new neutrino telescopes (AMANDA, Baikal, NESTOR, DUMAND) being planned or under construction (see the talks of these groups in these proceedings; an overview can be found in [17]).

The prediction of rates is quite involved, as one has to compute capture rates, fragmentation functions in basic annihilation processes, propagation through the solar or terrestial medium, charged current cross sections and muon propagation in the rock, ice or water surrounding the detector. In addition, there may be scattering of the Cherenkov photons generated by the muons, due to impurities in the medium.

The capture rate induced by scalar (spin-independent) interactions between the neutralinos and the nuclei in the interior of the Earth or Sun is the most difficult one to compute, since it depends sensitively on Higgs mass, form factors, and other poorly known quantities. For the Sun, the axial cross section is relatively easy to compute, a good approximation is given by [4]

$$C_{\rm ax}^{\odot} = (1.3 \times 10^{25}\,{\rm s}^{-1})\,\frac{\rho_{0.3}^{\chi}\,\sigma_{0\,spin}^{H(40)}}{(m_{\chi}/(1{\rm GeV}))\bar{v}_{270}} \qquad (14)$$

where $\sigma_{0\,spin}^{H(40)}$ is the cross section for neutralino-proton elastic scattering via the axial-vector interaction in units of $10^{-40}{\rm cm}^2$, $\bar{v}_{270}$ is the dark-matter velocity dispersion in units of 270 ${\rm kms}^{-1}$, and $\rho_{0.3}^{\chi}$ is the local halo mass density in units of 0.3 GeV ${\rm cm}^{-3}$. The capture rate in the Earth is dominated by scalar interactions, where there may be kinematic and other enhancements, in particular if the mass of the neutralino almost matches one of the heavy elements in the Earth. For this case, a more detailed analysis is called for, but convenient approximations are available [4].

Neutrinos produced by neutralino annihilations in the Sun may suffer energy loss one their way out through the solar medium. Charge-current reactions make neutrinos disappear whereas neutral current reactions degrade the energy of the original neutrino. Since these events are rare but "catastrophic" when they occur, they should be modelled event-by-event [18] instead of in a statistical fashion [19]. For the Earth such effects can be neglected in the energy range of interest.

To illustrate the potential of neutrino telescopes for discovery of dark matter through neutrinos from the Earth or Sun, we present the results of a full calculation [20]. In Fig. 2 it can be seen that a neutrino telescope of area around 1 ${\rm km}^2$, which is a size currently being discussed, would have discovery potential for the supersymmetric models with high scalar cross sections discussed above. It can also be seen that gamma line detection and neutrino detection are complementary techniques in the sense that the respective rates are not strongly correlated.

6 Conclusions

To conclude, we have seen that indirect detection methods have the potential to be very useful complements to direct detection of supersymmetric dark matter candidates. In particular, new

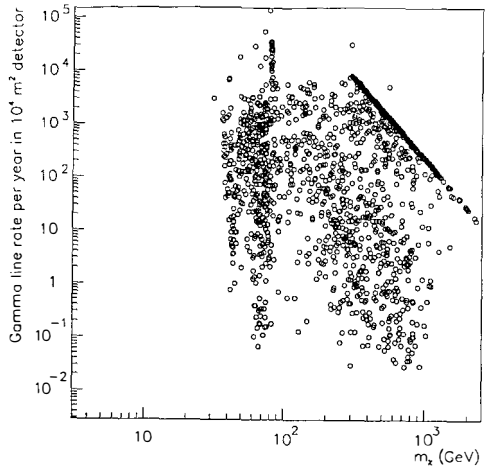


Figure 1: *The gamma line flux coming from* $\chi\chi \to \gamma\gamma$ vs *neutralino mass (= gamma ray energy) for a variety of supersymmetric models consistent with all accelerator bounds.*

air Cherenkov and neutrino telescopes may have the sensitivity to rule out or confirm the supersymmetry solution of the dark matter problem.

7 Acknowledgements

I wish to thank my collaborators, Joakim Edsjö and Paolo Gondolo for many helpful discussions. This work has been supported in part by the Swedish Natural Science Research Council (NFR) and the European Union Theoretical Astroparticle Physics Network (TAN).

References

1) N. Seiberg and E. Witten, Nucl. Phys. **B426** (1994) 19.

2) H.E. Haber and G.L. Kane, Phys. Rep. **117** (1985) 75.

3) U. Amaldi, W. de Boer and H. Furstenau, Phys. Lett. **B260** (1991) 447.

4) G. Jungman, M. Kamionkowski and K. Griest, Phys. Rep. to appear, hep-ph/9506380 (1995); G. Jungman, these proceedings.

5) L. Bergström and P. Gondolo, hep-ph/9510252 (Astroparticle Physics, in press); P. Gondolo, these proceedings.

6) F.M. Borzumati, M. Drees and M.M. Nojiri, Phys. Rev. **D51** (1995) 341.

7) M.S. Alam, et al. (CLEO Collaboration), Phys. Rev. Lett. **71** (1993) 674; Phys. Rev. Lett. **74** (1995) 2885.

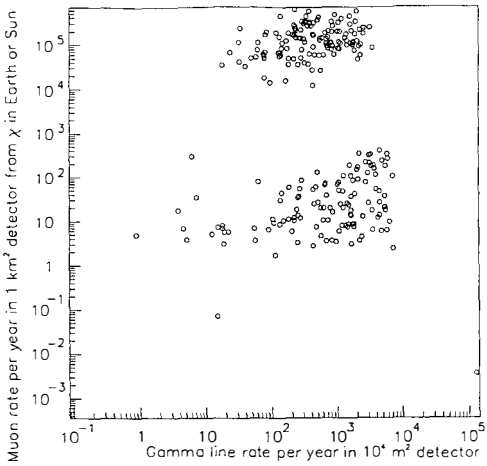


Figure 2: *The summed number of events per year of muons above 10 GeV in the direction from the centre of the Earth or the direction of the Sun expected for a neutrino telescope of size 1 km^2 vs the gamma line rate per year in a 10^4 m^2 air Cherenkov detector, for a range of supersymmetric models described in the text.*

8) S.W. Barwick et al., Phys. Rev. Lett. **75** (1995) 390.

9) J. Silk and M. Srednicki, Phys. Rev. Lett. **53** (1984) 624; F. Stecker, S. Rudaz and T. Walsh, Phys. Rev. Lett. **55** (1985) 2622. J. Ellis et al., Phys. Lett. **B214** (1989) 403; G. Jungman and M. Kamionkowski, Phys. Rev. **D49** (1994) 2316; A. Bottino, C. Favero, N. Fornengo, G. Mignola, Astropart. Phys. **3** (1995) 77.

10) S.P. Ahlen et al., Nucl. Instr. Methods **A350** (1994) 351.

11) J. Silk, K. Olive, and M. Srednicki, Phys. Rev. Lett. **55** (1985) 257; T. Gaisser, G. Steigman, and S. Tilav, Phys. Rev. **D34**, (1986) 2206; K. A. Olive and M. Srednicki, Phys. Lett. **B205** (1988) 553; L. M. Krauss, M. Srednicki, and F. Wilczek, Phys. Rev. **D33** (19 86) 2079; K. Freese, Phys. Lett. **B167** (1986) 295; K. Griest and D. Seckel, Nuc l. Phys. **B283** (1987) 681; F. Halzen, T. Stelzer, and M. Kamionkowski, Phys. R ev. **D45**, (1992) 4439; A. Bottino et. al, Astropart. Phys. **3** (1995) 65 ; G. F. Giudice and E. Roulet, Nucl. Phys. **B316** (1989) 429; G. Gelmini, P. Gondolo and E. Roulet, Nucl. Phys. **B351** (1991) 623; S. Ritz and D. Seckel, Nucl. Phys. **B304** (1988) 877.

12) L. Bergström and H. Snellman, Phys. Rev.**D37**, (1988) 3737; L. Bergström, Nucl. Phys. **B325**, (1989) 647; L. Bergström Phys. Lett. **B225** (1989) 372; S. Rudaz, Phys. Rev. **D39**, (1989) 3549; A. Bouquet,

P. Salati and J. Silk, Phys. Rev. **D40**, (1989) 3168.

13) L. Bergström and J. Kaplan, Astropart, Phys. **2** (1994) 261.

14) M. S. Turner, Phys. Rev. **D34** (1986) 1921.

15) G. Jungman and M. Kamionkowski, Phys. Rev. **D51** (1995) 3121.

16) M. Urban et al., Phys. Lett. **B293** (1992) 149.

17) For a review, see S.W. Barwick in Proc. of Trends in Astroparticle Physics, Stockholm, 1994 (L. Bergström, P. Carlson, P.O. Hulth and H. Snellman, eds.), Nucl. Phys. B (Proc. Suppl.) **43** (1995) 183.

18) J. Edsjö, hep-ph/9504205, University of Uppsala Diploma thesis; J. Edsjö, these proceedings.

19) K. Griest and D. Seckel, Nucl. Phys. **B283** 681 (1987) 681.

20) L. Bergström, J. Edsjö and P. Gondolo, to appear.

Détection indirecte de matiere caché supersymmetrique

Une revue est presenté sur les methodes differentes de detection de particules supersymmetriques neutres comprennant hypothetiquement le halo de notre galaxie.

Indirekt detektion av supersymmetrisk mörk materia

I denna artikel ges en översikt av olika indirekta metoder för detektion av den hypotetiska supersymmetriska meteria, som kan utgöra merparten av massan i Vintergatans halo.

WIMP ANNIHILATIONS IN THE SUN AND IN THE EARTH

J. Edsjö*
Department of Theoretical Physics, Uppsala University,
Box 803, S-751 08 Uppsala, Sweden

Abstract

Weakly-interacting massive particles (WIMPs) annihilating in the center of the Sun or the Earth may give rise to energetic neutrinos which might be discovered by astronomical neutrino detectors. The angular distribution of the neutrino-induced muons is considered in detail via Monte Carlo simulations. It is shown that large underground Čerenkov neutrino telescopes might be able to extract the WIMP mass from the width of the muon angular distribution.

1 Introduction

Weakly-interacting massive particles (WIMPs) with masses in the GeV–TeV range are among the leading non-baryonic candidates for the dark matter in our galactic halo. One of the most promising methods for the discovery of WIMPs in the halo is via observation of energetic neutrinos from annihilation of WIMPs in the Sun and/or the Earth [1].

Here we want to study if Čerenkov neutrino telescopes might realistically expect to be able to extract the WIMP mass from the muon angular distribution (once measured).

2 Annihilation channels and muon fluxes

WIMPs trapped in the core of the Sun and/or Earth can annihilate to a fermion-antifermion pair, to gauge bosons, Higgs bosons and gluons ($\chi\chi \to \ell^+\ell^-$, $q\bar{q}$, gg, $q\bar{q}g$, W^+W^-, Z^0Z^0, Z^0H^0, $W^\pm H^\mp$, H^0H^0). These annihilation products will hadronize and/or decay, eventually producing high energy muon neutrinos.

We have performed detailed Monte Carlo simulations of the hadronization and decay of the annihilation products using JETSET 7.4 [3], of the neutrino interactions on their way out of the Sun and of the charged-current neutrino interactions near the detector using PYTHIA 5.7 [3], and finally of the multiple Coulomb scattering of the muon on its way to the detector using distributions from the Particle Data Group [4].

*E-mail address: edsjo@teorfys.uu.se

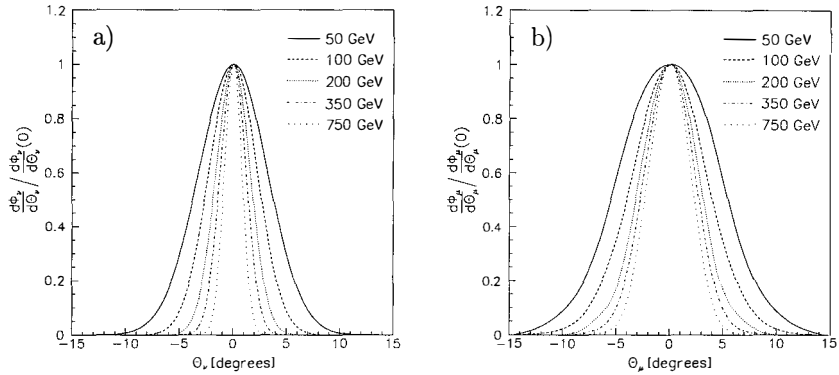


Figure 1: *Projected angular distributions of a) neutrinos and b) neutrino-induced muons from WIMP annihilations in the Earth. The distribution in b) are shown for hard channels (W^+W^- for 100–750 GeV and $\tau^+\tau^-$ for 50 GeV), with a detector muon threshold $E_\mu^{th} = 10$ GeV and a detector angular resolution $\theta_{det} = 1.4°$.*

With respect to previous calculations [5, 1], this Monte Carlo treatment of the neutrino propagation through the Sun bypasses simplifying assumptions previously made, namely neutral currents are no more assumed to be much weaker than charged currents and the energy loss is no more considered to be continuous. This results in that for a WIMP mass of 1500 GeV (50 GeV) we find a muon flux 20% (5%) higher than that obtained by Ritz and Seckel [5]. For details about the simulation, see Ref. [6, 2].

3 Annihilation profiles

The annihilation region in the Sun can be regarded as point-like, its angular size being $\lesssim 0.005°$ for the WIMP masses we are interested in, $m \gtrsim 10$ GeV. For the Earth, on the contrary, the angular extension of the annihilation region is non-negligible [1]. One can show [2] that the projected angular distribution from the Earth is approximately Gaussian with root mean square value $\theta_\nu^{rms} \simeq 23°/\sqrt{m/\mathrm{GeV}}$ for $m \gtrsim 10$ GeV.

Projected angular distributions of the neutrino flux from the Earth are shown in Fig. 1a for WIMP masses between 50 and 750 GeV. Note that these distributions are independent of the neutrino energy spectrum and of the specific annihilation channel. The analogous distributions for the Sun are simply narrow peaks at $\theta_\nu = 0$.

4 Muon angular distributions

In Čerenkov neutrino telescopes it is not possible to measure the angular distribution of the neutrinos directly since it is the muon produced in charged-current interactions that can be detected. The direction of the neutrino is somewhat lost because of two effects: (1) the muon produced in a charged-current interaction exits at an angle θ_{CC} with respect to the incoming neutrino and (2) the same muon undergoes multiple Coulomb scattering on its way to the detector, changing direction by an angle θ_{Coul}. We use the charged current angular distributions as given by PYTHIA and the multiple Coulomb scattering distributions given in Ref. [4]. Both angles get smaller with increasing neutrino (and muon) energy, their root mean square values

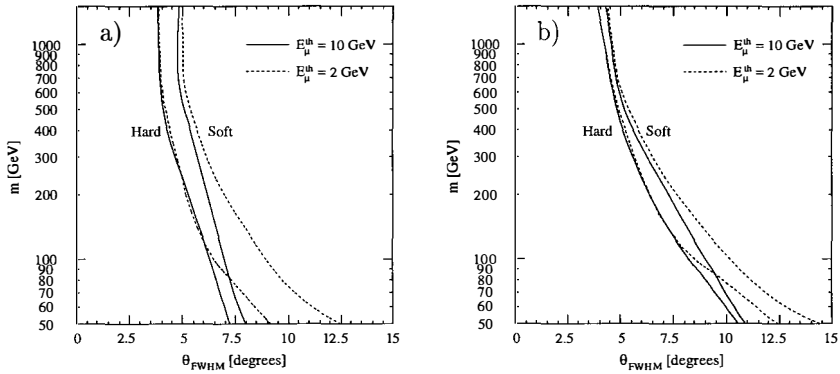


Figure 2: *WIMP mass m versus full width half maximum θ_{FWHM} of the neutrino-induced muon distribution for soft ($b\bar{b}$) and hard (W^+W^- and $\tau^+\tau^-$) channels. a) For WIMPs annihilating in the Sun with a detector angular resolution of $\theta_{det} = 1.4°$. b) For WIMPs annihilating in the Earth with a detector angular resolution of $\theta_{det} = 1.4°$.*

being approximately inversely proportional to the square root of the energy.

There is an additional uncertainty coming from the reconstruction of the muon track. Each neutrino telescope has an intrinsic angular resolution in determining the direction of the muon. We assume that the error in its determination is Gaussian with root mean square value θ_{det}, typically of the order of 1–2°.

In Fig. 1b we plot the muon angular distributions for hard channels in the Earth, where we by hard mean that it is dominated by a hard channel, like W^+W^-, Z^0Z^0 and $\tau^+\tau^-$. Since the rate is proportional to the neutrino energy squared[†] it suffices that the branching ratio into hard channels is greater than $\sim 10\%$. Softer neutrino spectra, *e.g.* those from the $b\bar{b}$ and H^0H^0 channels, give rise to wider angular distributions. This is due to the energy dependence of θ_{CC}^{rms} and θ_{Coul}^{rms}. Note the difference between the neutrino (Fig. 1a) and muon (Fig. 1b) angular distributions: charged-current interactions and multiple Coulomb scattering make the width dependence on WIMP mass stronger.

5 WIMP mass determination

In Figs. 2a-b we show the dependence of the WIMP mass on the full width half maximum of the projected muon angular distributions. We present the soft $b\bar{b}$-channel and the hard W^+W^-- and $\tau^+\tau^-$-channels for the Earth and the Sun for two different muon energy thresholds, $E_\mu^{th} =$ 2 GeV and $E_\mu^{th} = 10$ GeV.

We find that the detector angular resolution is the limiting factor for the mass determination of heavy WIMPs ($m \gtrsim 400$ GeV). For lighter WIMPs, it seems promising to infer their mass from the Earth muon distribution. We remind that this is also the mass range in which the signal from the Earth is expected to be significant. The WIMP mass could also be extracted from the Sun muon distributions provided the detector muon energy threshold is low. In fact, the width of the angular distribution for the Sun is dominated by charged-current scattering, which acts as a mass spectrometer in diffusing neutrinos according to their energies and so according to the WIMP mass.

[†]This is so since both the cross section and the muon range rises linearly with energy.

6 Conclusions

WIMPs annihilating in the center of the Sun or the Earth may give rise to a neutrino-induced muon flux in neutrino telescopes. The width of the muon angular distribution carries information on the WIMP mass, because the size of the annihilation region, the charged-current neutrino-nucleon scattering and the multiple Coulomb scattering of the muons all depend on the WIMP mass. It has been shown that the WIMP mass can be inferred, for WIMPs lighter than $\sim$ 400 GeV, from the Earth distribution and, provided the muon energy threshold is low ($\lesssim$ 5 GeV), also from the Sun distribution. This seems therefore a promising method of determining the WIMP mass.

Acknowledgments

The work on the angular distributions has been performed in collaboration with P. Gondolo. This work has been partially supported by the EC Theoretical Astroparticle Network under contract No. CHRX-CT93-0120 (Direction Générale 12 COMA).

References

1) See Ref. [2] for a detailed list of references.

2) J. Edsjö and P. Gondolo, Phys. Lett. **B357** (1995) 595.

3) T. Sjöstrand, Comp. Phys. Comm. **82** (1994) 74. T. Sjöstrand, *PYTHIA 5.7 and JETSET 7.4. Physics and Manual*, CERN-TH.7112/93, hep-ph/9508391 (revised version).

4) Particle Data Group, Phys. Rev. **D50** (1994) 1173.

5) S. Ritz and D. Seckel, Nucl. Phys. **B304** (1988) 877.

6) J. Edsjö, Diploma Thesis, Uppsala University preprint TSL/ISV-93-0091 (ISSN 0284-2769). J. Edsjö, Nucl. Phys. **B** (Proc. Suppl.) 43 (1995) 265.

Annihilations des WIMPs dans le Soleil et dans la Terre

Résumé

WIMPs annihilant dans le centre du Soleil ou de la Terre peuvent produire neutrinos energique qui on peut découvrir avec des détecteurs astronomiques des neutrinos. La distribution angulaire des muons (de l'interaction neutrino-nucleon) est examiné par des simulations Monte Carlo. Nous montrons que des téléscopes des neutrinos peuvent extraire la masse de la WIMP par la largeur de la distribution angulaire des muons.

THE NESTOR EXPERIMENT

Yves SACQUIN*

CEA Saclay, DAPNIA/SPP
F-91191 Gif/Yvette CEDEX

Abstract

The NESTOR collaboration intends to build a high energy neutrino detector located on deep sea bottom. The physics topics of the projects include astrophysics, neutrino physics and dark matter search in the form of heavy neutralinos. The expected performances and the present status are described. The major works under way are reviewed, together with the essential milestones.
The NESTOR collaboration participates to the discussions in view of a future international one km^3 detector.

1 Introduction

NESTOR is an experiment dedicated to the search of high-energy astrophysical neutrinos. It is based on the detection of the Tcherenkov light radiated by the muon issued from a muon neutrino interaction, in a transparent medium like the sea water; the detector needs to be shielded from the muons produced by the cosmic ray interactions in the atmosphere, therefore it is located on the deepest possible submarine ground. The interactions of neutrinos in the sea water or in the sublying rocks can therefore be distinguished more easily from the background muons coming from atmospheric sources.

Benefiting from the DUMAND experience, NESTOR aims to deploy, in a single procedure, a tower structure already large enough to reach many physics goals [1]. The full physics program needs the deployment of seven towers, a first step towards a one km^3 detector.

The physics program includes the study of point-like or diffuse neutrino sources – galactic or extragalactic – like the active galactic nuclei (AGN) or the supernovae remnants, and the study of neutrino oscillations using atmospheric neutrinos induced by high energy cosmic rays

[1]http://infodan.in2p3.fr/delphi/user/sacquin/sacquin.html

interacting with the terrestrial atmosphere, or accelerator produced neutrinos. The detection of WIMPs in the form of supersymetric neutralinos will also be looked for: we refer the reader to the contributions of J. Edsjō and of L. Bergström, in the same proceedings, for the production modes and expected rates. Unexpected signals will also be looked for, owing to the fact that a full transmission of all data is foreseen.

The collaboration presently consists of high energy physicists, astrophysicists and oceanologists from 18 institutes in 6 countries (Greece, Russia, France, Germany, Italy, U.S.A.).

2 Apparatus

The NESTOR structure (Figure 1) consists in a semi-rigid buoyant tower made of 12 floors, anchored on the sea bottom. Each floor is made of 6 rigid titanium arms, 16 m long, whose extremities are equipped with optical modules. The arms are fixed to the center axis of the tower, and the center of each floor is also equipped with optical modules. Two adjacent floors are vertically 20 m from each other.

An optical module consists in a 15" Hamamatsu phototube, housed in a 17" oceanological glass sphere (Benthosphere). There are 7 pairs of optical modules on each floor, each pair having one PM looking downwards and one looking upwards. The optical modules are connected to a titanium sphere in the middle of each floor, housing the electronics for the signals and the slow control. From each floor, an optical fiber transmits the signals to a submarine opto-electrical cable, with 12 optical fibers, going to the shore station.

The transmission of the signals will be either digital or analog, depending of the performances reached in the presently ongoing developpments of both options. All signals above a threshold of one quarter of photoelectron will be transmitted, and used for a physics trigger by the acquisition station on the shore.

The slow control includes a calibration system with light emitting diodes in glass sphere in between two floors, an accurate sonar positioning system and a continuous monitoring of environmental parameters (temperatures, currents, magnetic field, tilt measurements...).

3 Environment

The chosen site of NESTOR is located in the Ionian Sea near the city of Pylos, South-West of Peloponesos (Greece). There lies a flat submarine plateau at 3800m depth, 11 nautical miles away from the shore, on a wide area (5 nautical mile large), which is convenient for the future extension to 7 towers.

Several cruises with oceanological vessels have been dedicated to the environmental studies and hydrographic survey of the site. The water transmissivity to light has been monitored, and found to be of the order of 55m for the wavelength of the Tcherenkov light. Underwater currents have been measured in several cruises, at values less than 5 cm/s. The temperature of the water at the bottom is around 14^0C, and the water salinity and radioactivity has been measured at an acceptable level (13 Bq/l).

4 Performances

The reconstruction of the muon trajectory parameters is the result of a minimization using the time and amplitude recorded on each module. This ensures an angular resolution from 1 to 3 degrees depending on the azimuth and distance of the track. The effective area covered by one tower varies with azimuth from 10 000 to 18 000 m^2. A 4π acceptance is covered, and muons with energies as low as 5 GeV can be reconstructed. An approximate energy determination is possible, using the amplitudes of the signals.

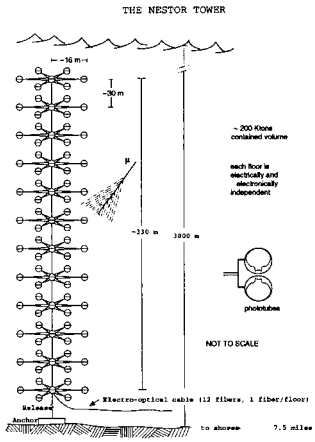


Figure 1: Sketch of the NESTOR tower with its 12 floors, and a close-up view of a pair of optical modules, made of two spheres with PM looking up and down.

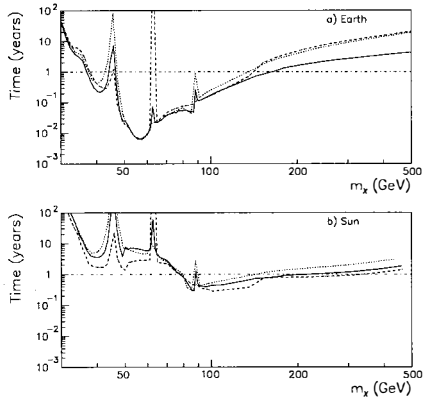


Figure 2: Minimal exposure time necessary to detect a 4σ effect in neutralino search, with one NESTOR tower. a) refers to a signal coming from the Earth core and b) for neutrinos coming from the Sun. The three different curves refer to different gaugino fractionnal weight P: dotted line is for $P = 0.1$, solid line for $P = 0.5$, dashed line for $P = 0.9$.

For the point-like sources search, a rate of a few events per year is expected using standard prediction [2]. The diffuse AGN background should provide a few tens to a few hundreds of events per year, still with one tower. The contamination due to atmospheric neutrinos is important only below 1 TeV.

For the neutralino search [3], the knowledge of the exact direction of the muon (the Earth core or the Sun) allows a better rejection of the background and a better constraint on the fit. As seen in figure 2, one year of exposure time of one tower allows to put a limit on the neutralino mass at around 150 GeV/c^2. A longer exposure time can extend greatly this limit, specially looking at the signal from the Sun.

The study of neutrinos from atmospheric origin allows to look for neutrino oscillations, with very long base-line. The surveyed domain is comparable to the one covered by the Kamioka experiment, which has claimed a positive result.

Eventually, a neutrino beam coming from CERN is under study. Such a beam would allow to investigate a large area of the $\sin^2 2\theta$-Δm^2 plane.

5 Present studies in mechanics and electronics

Presently the collaboration is actively working on the definition of the mechanical structure and on the transmission scheme.

The original drawing of the structure results from a titanium smaller structure deployed in 1992 by the russian collaborators. They are presently building a full size two floors prototype, which should be delivered in spring and tested on the NESTOR site. The deployment procedure, which is crucial for this kind of apparatus, should lead to a better definition of the tower.

In the mean time, engineers from Saclay and Marseilles, an active observator yet in

NESTOR, are exploring alternative options, whith the help of french oceanological companies. A floor mock-up is tested in wind tunnel, and studies continue in oceanological research basin. This should lead to a white paper on the deployment and structures before summer.

Two alternative transmission schemes are explored in parallel: A digital solution is developped by the Rome group, now building a two channels prototype, to be used in the forecoming tests. An important parameter of this option, the MTBF (mean time before failure) is presently under study.

An analog transmission of the signals is studied in Saclay. A three channel prototype is about to be delivered, and tested with a long optical fiber. The analog signals are multiplexed on a radio frequency which modulates a laser source. A hybrid solution with an additionnal laser multiplexing is also investigated.

6 Summary and prospects

The necessary material for the optical modules of one tower is already bought and tested in Athens. The construction of mu-metal cages is in progress.

Work is progressing on the trigger, developped on DecPeRLe, on the shore station and laboratories, and on the opto-electical cable deployment.

Deployments of strings with one or two optical modules is scheduled this year, to extend the study of the light background conditions in the site, of the sedimentation and currents, and of the sea-water transparency.

By the end of 1996, the immersion of two floors and the validation of a transmission scheme shall be the starting point of the complete definition of the apparatus. The building of a tower is planned to follow during the next two years. The project will afterwards be completed by extending to seven towers.

Eventually a one km^3 detector is foreseen, for which a world-wide collaboration is in active discussions.

References

[1] L.K. Resvanis, editor, Proceedings of the 2nd NESTOR international workshop, Pylos, Greece, 1992.
L.K. Resvanis, editor, Proceedings of the 3rd NESTOR international workshop, Pylos, Greece, 1993.

[2] See ref. [1] and references in NESTOR proposal, volume 1, in preparation (The author can provide them until it is issued).

[3] V. Berezinsky, A. Bottino, J. Ellis, N. Fornengo, G. Mignola and S. Scopel, "Neutralino dark matter in supersymmetric models with non-universal scalar mass terms.", preprint CERN-TH 95-206, August 1995.

Résumé

La collaboration NESTOR a pour but de construire un détecteur de neutrinos de haute énergie, installé en mer à grande profondeur. Les objectifs de physique couverts par le projet concernent l'astrophysique, la physique des neutrinos, ainsi que la recherche de matière noire sous la forme de neutralinos massifs. Les performances attendues du détecteur ainsi que l'état actuel du projet sont présentés. Les principaux travaux en cours sont passés en revue, ainsi que les étapes essentielles du projet.
La collaboration NESTOR entend participer à la réalisation d'un futur détecteur mondial de un km^3.

Status of the Projects BAIKAL and AMANDA

CH. SPIERING
DESY-Zeuthen, Platanenallee 6, D-15738 Zeuthen, Germany

1 Introduction

The idea to detect the Cherenkov light emitted by particles produced in neutrino interactions in big water basins was proposed in 1960 [1]. In spite of the efforts of the pioneering experiments DUMAND (in the Pacific near Hawaii) and BAIKAL (in the Siberian Lake Baikal), it was not until 1993 before the first small underwater telescope went into operation. This was *NT-36* ("Neutrino Telescope with 36 phototubes"), a detector in Lake Baikal carrying 36 PMTs at 3 strings. After several upgrades, a 96-PMT telescope takes data since April 1996.

The BAIKAL project aims to build a large deep underwater detector for muons and neutrinos in a depth of about 1 km. The specifics of the project is the ice cover of the lake which allows for easy deployment and retrieval of detector components in late winter. This advantage could balance the drawbacks of the site: the comparatively low depth, giving a high background rate of atmospheric muons punching through the water to 1 km depth; and the moderate absorbtion length of about 20 m (compared to 40-50 m in Oceans and more than 100 m in the Antarctic ice).

AMANDA is a project to build a large neutrino telescope in the deep ice at the South Pole. The motivation to move to a remote location like the Pole has been given by several arguments: *a)* the geographical location is complementary to the other detectors, *b)* ice is a sterile medium with a background noise determined only by the intrinsic dark noise of a PMT, and *c)* the ice provides a stable deployment platform, the position of OMs does not change once they are placed, and the distance between detector and DAQ system is kept short, 1-2 km compared to 7-30 km in the deep water experiments. The optical properties of the ice are different compared to those of Lake Baikal. Whereas the absorption length is extremely large, the influence of scattering is much higher than in water.

In 1994, the AMANDA group deployed a 80-PMT array in 0.8-1.0 km depth. Residual bubbles in the ice turned out to cause unacceptably high light scattering in the ice. Therefore, in 1996 a detector of 86 PMTs was deployed at depths ranging from 1500 to 2000 meters – with scattering reduced by about two orders of magnitude.

The present report describes the status of the two projects.

2 Status of the Lake BAIKAL Telescope

The BAIKAL Collaboration: I.A.Belolaptikov[7], L.B.Bezrukov[1], B.A.Borisovets[1], N.M.Budnev[2], A.G.Chensky[2], I.A.Danilchenko[1], J.-A.M.Djilkibaev[1], V.I.Dobrynin[2], G.V.Domogatsky[1], A.A.Doroshenko[1], S.V.Fialkovsky[4], A.A.Garus[1], A.Gaponenko[2], O.A.Gress[2], T.A.Gress[2], H.Heukenkamp[8*], S.B.Ignatev[3], A.Karle[8], A.M.Klabukov[1], A.I.Klimov[6], S.I.Klimushin[1], A.P.Koshechkin[2], J.Krabi[8], V.F.Kulepov[4], L.A.Kuzmichov[3], B.K.Lubsandorzhiev[1], M.B.Milenin[4], T.Mikolajski[8], R.R.Mirgazov[2], A.V.Moroz[2], N.I.Moseiko[3], S.A.Nikiforov[2], E.A.Osipova[3], A.I.Panfilov[1], Yu.V.Parfenov[2], A.A.Pavlov[2], D.P.Petukhov[1], K.A.Pocheikin[2], P.G.Pokhil[1], P.A.Pokolev[2], M.I.Rosanov[5], V.Yu.Rubzov[2], S.I.Sinegovsky[2], I.A.Sokalski[1], Ch.Spiering[8], O.Streicher[8], P.A.Pokolev[2], T.Thon[8], I.I.Trofimenko[1], Ch.Wiebusch[8], R.Wischnewski[8]

[1] *Institute for Nuclear Research, Russian Acad. of Science (Moscow, Russia),* [2] *Irkutsk State University (Irkutsk, Russia),* [3] *Moscow State University (Moscow, Russia),* [4] *Nizhni Novgorod State Technical University (Nizhni Novgorod, Russia),* [5] *St.Petersburg State Marine Technical University (St.Petersburg, Russia),* [6] *Kurchatov Institute (Moscow, Russia),* [7] *Joint Institute for Nuclear Research (Dubna, Russia),* [8] *DESY Institute for High Energy Physics (Zeuthen, Germany)*

2.1 Detector and Performance

The Baikal Neutrino Telescope [2)] is being deployed in the Siberian Lake Baikal, at a depth of 1.1 km, see fig.1. At April 13th 1993, the detector *NT-36* (36 PMTs at 3 strings) was put into operation and took data up to March 1994. During 240 days of data taking, $7 \cdot 10^7$ events for the basic trigger 3/1 ($\geq$ 3 hits at $\geq$ 1 string) have been taken, 10^7 of them fulfilling the trigger condition 6/3 (i.e. $\geq$ 6 hits at 3 strings), suitable for unambiguous track reconstruction. A modified array, *NT-36'*, came into operation at April 3th, 1994. During 242 days operating time, $9.7 \cdot 10^7$ events for trigger 3/1 have been taken. An array carrying 72 PMTs has been deployed in March 1995 (see fig.2). In April 1996, a 96-PMT variant has started data taking. These arrays are steps towards the Neutrino Telescope *NT-200* which will carry 192 PMTs.

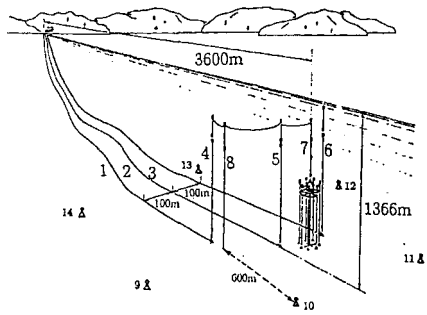


Figure 1: Installations at the Baikal site (status 1995); 1-3: shore cables, 4-6: string stations for shore cables, 7: string with the telescope, 8: environmental string, 9-14: ultrasonic emitters.

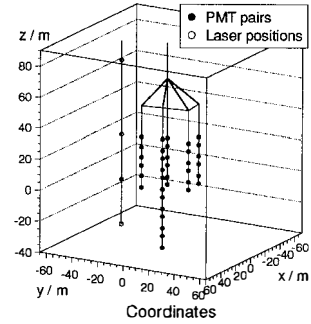


Figure 2: Schematical view of NT-72, deployed in 1995. It consists of 36 PMT-*pairs* at 4 short and 1 long string. Open circles indicate positions of the distant calibration laser (see text).

The 7 arms of the frame carrying the detector, each 21.5 m in length, are at a height of 250 m above the bottom of the lake. Two underwater electrical cables and one optical cable connect the detector site with the shore station. Detector components are deployed form the ice cover in later winter.

The optical modules are grouped in pairs along the strings, directed alternatively upward and downward. The two PMTs of a pair are connected in coincidence and define a *channel.* The local coincidences are mandatory for the suppression of the background from bioluminescence and PMT noise: the typical 1 *p.e.* counting rate is 100 kHz for a single tube, but only 200 Hz for a channel.

A *muon-trigger* is formed by the requirement of $\geq 3(4)$ fired channels within a time window of 500 nsec. A second system (*monopole trigger*) searches for time patterns characteristic for slowly moving, bright particles like nuclearities or GUT magnetic monopoles catalyzing proton decays.

The time calibration of the array is performed with the help of a stationary nitrogen laser positioned just above the array. The spatial position of the components of the array is monitored by a hydroacoustical system. A special "environmental" string carries devices to measure the optical parameters of the water, temperature, pressure, sound velocity and water currents.

The performance of the detector can be summarized as follows:

- Seasonal variations of the water luminescence (reaching sometimes an order of magnitude) do not influence the muon trigger rate.
- Upward pointing modules loose after 150 days $\approx$50 % of their sensitivity due to sedimentation, downward modules do not show any remarkable change of their sensitivity. That means that for long term operation, the optical modules dominantly should point downwards. We are searching for methods to reduce the settlement of matter on the glass spheres.
- Displacements of the whole array due to water currents are small (< 2m) and monitored by our ultrasonic system. The relative positions of the optical modules do not change by more than 20 cm.
- All mechanical components of the system worked reliably, particularly none of the modules did leak over the first 2 years. On the other hand, some of the electronical components do not satisfy the standards requested for long-time remote operation.

2.2 Selected results

First physical results based on the data taken with the 1993 array have been presented at different occasions (see ref.3 and refs. therein). Some of the results are given in the following:

a) Muon Spectra

Muon angular distributions are well described by MC expectations. Converting the measured angular dependence into a depth dependence of the vertical flux, good agreement with other published values is observed (fig.3).

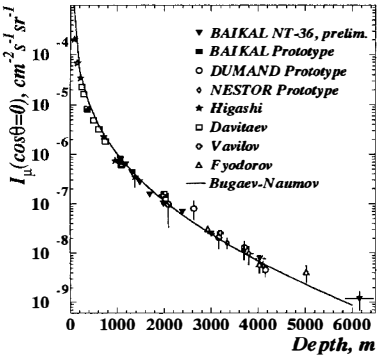


Figure 3: Vertical muon flux vs. water depth.

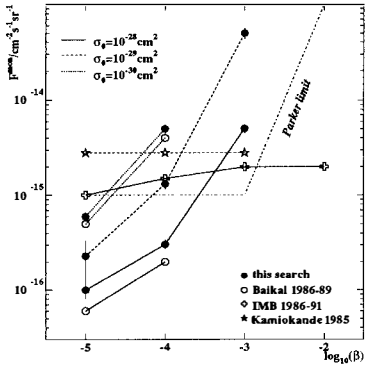


Figure 4: Upper limits (90% CL) on the flux of magnetic GUT monopoles as a function of their velocity β, for different catalysis cross sections σ_o.

b) Search for Neutrino Candidates

With *NT-36*, a signal-to-noise ratio, S/N of 1:50 has been achieved over the full lower hemisphere. Here, the signal is given by upward muons from interactions of atmospheric neutrinos, and the noise are downward muons faking upward muons.

In selected angular regions, however, S/N may be much larger. Searching for muons just around the opposite zenith (as expected from neutralino annihilation in the center of the Earth) even with the samll *NT-36*, the S/N-ratio approaches unity. A first promising candidate for a nearly vertical upward muon was found in the 1994 data. S/N for this events is estimated to be about 3, i.e. the probability to be a fake event is only 25%.

c) Search for Magnetic GUT Monopoles

From monopole induced proton decays one would expect Cherenkov light signals generated by the decay products. For certain regions of the parameter space in β (monopole velocity) and σ_c (catalysis cross section), GUT monopoles would cause sequential hits in individual channels in time windows of 10^{-4} - 10^{-3} sec. Having searched for such enhanced counting rates we deduce upper limits for the flux of monopoles catalyzing the decay of free protons with cross section $\sigma_c = 0.17 \cdot \sigma_o \cdot \beta^{-2}$. E.g., with $\sigma_o = 10^{-29}$ cm^2 and $\beta = 10^{-4}$ one gets a flux limit of 10^{-15} cm^{-2} sec^{-1} sr^{-1}, just at the Parker limit and of the same order of magnitude as limits obtained by our collaboration with dedicated setups in 1984-89 (fig.4).

d) Laser Experiment 1995

Just after deployment of NT-72, a calibration experiment with a *distant* laser was performed. Fig. 1 indicates the different positions of the laser. An isotropic light source was used to simulate distant electromagnetic showers. From the time and amplitude pattern of the hit channels light attenuation and scattering lengths, λ_a and λ_s, could be determined. We found $\lambda_a = 22.7\,m$ at $475nm$.

2.3 Future plans

The main goal of 1996 is to identify neutrinos with the help of NT-96. S/N for this array was calculated to be larger than 1 over most of the lower hemisphere, with 40-60 neutrino events to be taken with a fully functioning detector over one year. For 1997, an extended array with a series of qualitative improvements is scheduled. NT-200 is planned to be finished in spring 1998.

3 Status of the AMANDA Project

The AMANDA Collaboration: P. Askebjer [1], S.W. Barwick [2], L. Bergström [3], A. Bouchta [1], S. Carius [3], E. Dahlberg [1], B. Erlandsson [1], A. Goobar [1], L. Gray [4], A. Hallgren [3], F. Halzen [4], H. Heukenkamp [5], P.O. Hulth [1], St. Hundertmark [5], J. Jacobsen [4], S. Johansson [1], V. Kandhadai [4], A. Karle [5], I. Liubarsky [4], D. Lowder [6], T.C. Miller [7], T.Mikolajski [5], P.C. Mock [2], R. Morse [4], R. Porrata [2], P.B. Price [6], A. Richards [6], H. Rubinstein [3], C. Spiering [5], O. Streicher [5], Q. Sun [1], T. Thon [5], S. Tilav [4], C. Walck [1], R. Wischnewski [5], G. Yodh [2]

[1] *Stockholm University, Stockholm, Sweden,* [2] *University of California, Irvine, USA,* [3] *University of Uppsala, Uppsala, Sweden,* [4] *University of Wisconsin, Madison, USA,* [5] *DESY - Inst.High En. Physics, Zeuthen, Germany,* [6] *University of California, Berkeley, USA,* [7] *Bartol Research Institute, University of Delaware, Newark, USA*

Rather than in water, the optical modules of AMANDA (Antarctic Muon And Neutrino Detector Array) are deployed in the 3 km thick ice cover at the geographical South Pole [4]. Holes are drilled in the ice with the help of a hot water drill. The strings with the optical modules are deployed into the water column in the hole which then refreezes within some days.

The only elements under ice are 8" PMTs (including HV dividers) in pressure spheres. Each PMT has its own cable supplying the HV and transmitting the anode signal in the other direction. All electronics is housed in a comfortable building at the surface. The simplicity of the components under ice and the non-hierarchical structure makes the detector highly reliable.

A muon trigger is formed by the requirement of a minimum number of fired PMTs within a time window of 2μsec. A second system searches for enhanced counting rates over time windows of some seconds (which could be due to the vast amount of low energy neutrino reactions induced by a supernova collapse). The same system also looks for time patterns which might be due to slowly moving bright objects.

AMANDA profits from the excellent logistics and infrastructure created to support scientific research at the Pole.

3.1 AMANDA-*A*

Preliminary explorations of the site and the drilling technology were performed in the antarctic summer 1991/92. During the 1993/94 campaign, four strings each with 20 PMTs have been frozen between 800 and 1000 m depth (AMANDA-*A*, see fig.6). Along with the coax cable, optical fibers carry light from a laser at the surface to each PMT. Each fiber was terminated by a light diffuser placed about 30 cm from the PMT. By sending laser pulses to individual diffusers and measuring the photon arrival time distributions at *distant* PMTs, the optical properties of the medium could be derived. The timing distributions indicated that photons do not propagate

straight-forward but are scattered and considerably delayed due to residual bubbles in the ice. The scattering prevents a proper reconstruction of the Cherenkov cone. One year later these measurements have been repeated in more detail and at different wavelengths. It turns out, that ice has an enormous absorption length, exceeding 200 m at a wavelength of 400 nm! The scattering length increases with depth, from 10 cm at 800 m depth to 20 cm at 1000 m depth (fig.7), with 20 cm scattering length of bubbles corresponding to an *effective* scattering length of about 70 cm [5] . In accordance with measurements at Vostok (East Antarctica) and Byrd Station (West Antarctica) these results suggest, that at depths greater than 1300-1400 m the bubbles will have disappeared and converted into clathrate crystals [6].

AMANDA-*A* is used for supernova search (see below), and to study events with high energy deposition and events registrated in coincidence with the South Pole surface array SPASE [6].

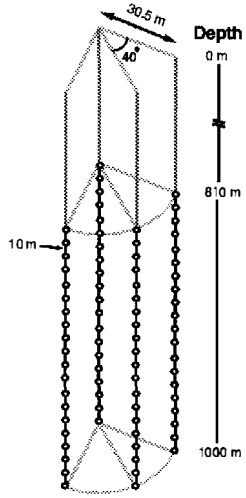


Figure 5: Schematical view of the shallow array AMANDA-*A*

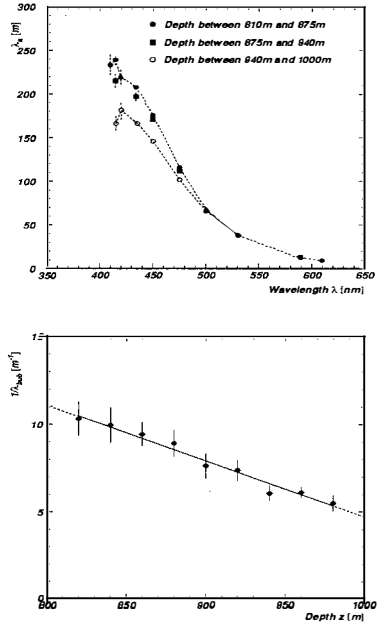


Figure 6: AMANDA-*A*: *Top* – Absorption length vs. wavelength, *bottom* – inverse scattering length vs. depth.

3.2 AMANDA-*B*

In January 1996, strings were deployed in greater depth. The main aim of the 1995/96 campaign was the exploration of the ice parameters in greater depth and the test wether the deep ice allows to construct a neutrino telescope.

Four strings have been deployed at depths between 1520 and 2000 m. The strings were equipped with 8-inch PMTs from Hamamatsu (R5912-2). With 14 dynodes and an amplification of 10^9 these tubes are tailored to drive signals over 2 km coax cable. Each string carries 20 PMTs with a vertical spacing of 20 m. The fourth string, in addition, carries 6 PMTs below the "standard" ones, whose signals are transmitted via twisted pair cable. The depths of the strings are 1520-1900 m (strings 1 and 2), 1570-1950 m (string 3) and 1570-1930 m ("standard" PMTs of string 4). The lowest of the PMTs equipped with twisted pair cables (string 4) is at 2000 m depth. All PMTs look down with the exception of levels 1 and 10 (string 1-3) and levels 1,2,10,19,20 (string 4). String 1-3 form a triangle with side lengths 84-57-52 m, string 4 is near the center (fig.7).

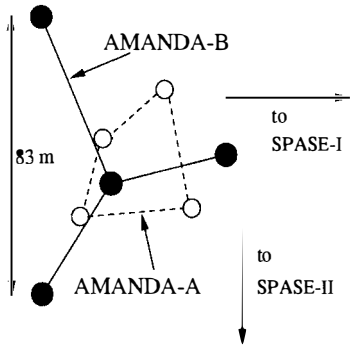


Figure 7: Top view of AMANDA-*B* and AMANDA-*A*

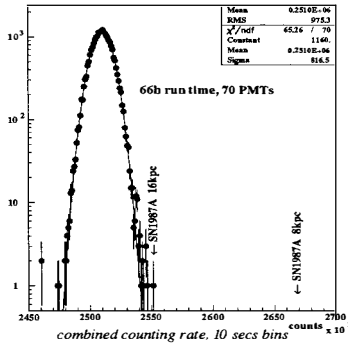


Figure 8: Distribution of the number of hits per 10 sec, summed over 70 PMTs of AMANDA-*B*.

The deployment procedure is well under control. The typical drilling time for 2 km depth is 3.5 days. The diameter of a hole is 50 cm. The deployment of one string took about 18 hours (with a limit of 35-40 hours set by the refreezing hole). Seven of the 86 PMTs were lost during the refreezing period, the remaining work stable. The time resolution for 1-PE signals was measured to be better than 5 nsec.

Various light calibration sources allow investigation of the optical properties of the ice:

- Optical fibers with diffusers for each OM. Fed by tunable laser at surface. Range of transmitted wavelengths $\geq$ 450 nm, time resolution about 15 nsec at 530 nm.
- Nitrogen laser at 1850 m depth, wavelength 337 nm, time resolution 1 nsec.
- Three DC halogen lamps (one broadband and two with filters for 350 and 380 nm).
- Six LEDs (operated in both modes, pulsed and DC), wavelength 390 nm and 450 nm.

The pulsed sources allow the determination of both absorption and scattering lengths by measuring timing distributions and the decrease of amplitudes with increasing distances. DC sources allow determination of the attenuation length by measuring the counting rate as a function of the distance. A *natural* calibration source is given by cosmic muons.

Preliminary analyses indicate, that for short wavelengths the absorption length is larger than 100 m. The effective scattering length seems to be about hundred times larger than at the depth of 800 m.

3.3 Physics Capabilities and Plans for the Next Season

The main goal of AMANDA-B is the investigation of ice properties. Hopefully, first upward going muons from neutrino interactions will be separated, including a few neutrino events identified by coincidences between AMANDA-B and AMANDA-A.

Energetic cascades will be searched for in both arrays. For the shallow array, the fast isotropization of light results in a calorimeter-like mode, with efficient light collection but no directional information of the showers. The deep array will allow crude pointing for cascades.

Coincidences between the surface air shower detector SPASE and AMANDA will be used to investigate muons in cosmic ray showers and to calibrate AMANDA.

We continue to search for counting rate excesses related to supernovae (time scales of seconds), Gamma Ray Bursts (time scales of msec) and non-relativistic bright particles (scales of hundred μsec). Fig.8 shows the distribution of the number of hits, summed over 70 PMTs of AMANDA-B, for adjacent 10 sec time intervals and a total run time of 66 hours. Arrows indicate the rate expected rate excess for a supernova at 8 and 16 kpc distance, respectively. Obviously, a supernova in most parts of our galaxy can be detected with great statistical certainty.

AMANDA-B will be upgraded next season with 6-7 new strings. Most of them will be equipped with twisted pair cables. However, a significant fraction of the PMTs will be used to test alternative technologies like analog transmission of the anode signal via optical fibers, or transmission of digitized data.

References

[1] M.A.Markov, Proc. 1960 Conf. on High Energy Physics, Rochester (1960) 578.
[2] I.A.Belolaptikov et al., Nucl.Phys.B (Proc.Suppl.),19 (1991) 375.
[3] Baikal collab., Proc.24rd ICRC, (Rome 1995),Vol.2, 536,742,770,789,841,1001,1043.
[4] S.Tilav et al., Proc. 23rd ICRC (Calgary 1993), vol.4,page 561.
[5] P.Askjeber et al., Science 267 (1995) 1147.
[6] Amanda collab., Proc.24rd ICRC, (Rome 1995), Vol.2, 758,777,816,1009,1011,1039.

L'etat des projets BAIKAL et AMANDA

Nous resumons l'etat actuel des telescopes de neutrino sous construction au lac Baikal en Siberie et au pole de Sud. Les resultats de l'experience Baikal comprennent un premier candidate de neutrino et des limites superieures de monopoles magnetiques. En 1996 le groupe Amanda a developpe un detecteur dans une profondeur de 1500 a 2000 m. Nous resumons la configuration et les capabilites du reseau du detecteur.

CAN ATMOSPHERIC CERENKOV TELESCOPES RESOLVE THE DARK MATTER PROBLEM?

Eric Paré
LPNHE, Ecole Polytechnique
91128 Palaiseau, France

Abstract

As a probe of the Universe through its most energetic gamma production, the Atmospheric Cerenkov Technique (ACT) could contribute to resolve the dark matter problem. Two such possibilities are examined. WIMPs may be observed through the gammas produced by their pair annihilations. Probing how the gamma from AGN's are absorbed by the light the galaxies deposed into the intergalactic medium may yield information on galaxy formation. The state of the art in ACT is outlined and the development of this activity, in France and elsewhere, is described.

1. Gamma ray astronomy above 10 GeV and the dark matter problem:

Above 10 GeV, the gamma ray fluxes expected are so small that only ground-based telescopes are viable. Quite recently, with the detection of the Crab nebula by the Whipple obs.[1], it has been demonstrated that the atmospheric Cerenkov technique is sensitive enough to open a new window in the electromagnetic spectrum around 1 TeV. Since then, another nebula[2] and two extragalactic sources (AGN) have been detected[3, 4] . Second generation telescopes, with higher sensitivity, are now under construction. Several projects, which will be in operation before the end of the century, will push the energy threshold down to the 10 GeV domain and the sensitivity down to the limit of the technique. These new instruments can contribute to the dark matter problem either by detecting gammas coming from the WIMP annihilation or by observing gamma ray absorption that will give information on the epoch of the galactic formation.

WIMP annihilation

A popular cold dark matter candidate is the lightest neutral supersymmetric particle called the neutralino. The mass of this particle is above 18.4 GeV from accelerator searches. In the minimum supersymmetric model, the neutralino is a mixture of Higgsino and Bino depending on the values of a few free parameters (μ, $\tan\beta$,..) [5]. Given the values of these parameters, annihilation cross sections can be calculated and the relic density estimated. The mass of the neutralino should be below 3 TeV

(pure Higgsino case) or below 600 GeV (pure Bino) in order not to overclose the universe ($\Omega < 1$). The annihilation process would then be a source of gamma rays above 10 GeV in the present day Universe. Gamma rays can appear in the decay chain of primary annihilation products (quarks, W, ...) and their energy will be continuously distributed. Alternatively, they can appear as lines at the neutralino mass (or a bit less) in two-body annihilation processes like $\gamma\gamma$ or γZ°, the branching ratios being of the order of a percent [6]. It may be difficult to prove that the soft gammas are not emitted by an ordinary astrophysical object, but detection of a line would be absolutely conclusive as no astrophysical process is known that emits lines in this energy range. The mean energy of the soft component is very low compared with the neutralino mass, so very low energy threshold detectors are needed. For the line detection, good energy resolution is necessary.

The emission rate depends on calculable annihilation cross sections and on the density (squared) of the neutralino. If it turns out that the neutralinos are distributed in the Galaxy as in the standard isothermal model, then the fluxes are much too low to be detected by Cerenkov detectors. According to Ipser and Sikivie[7], however, the dark matter density may be enhanced by a factor near 1000 at the Galactic centre. The emission, concentrated 150 pc from the centre, would appear as a diffuse source ($\Delta\Omega \approx 10^{-3}$ str). The gamma ray fluxes are expected to be around 10^{-13} cm^{-2} s^{-1} [8] but can vary by two orders of magnitude depending on the parameters of the supersymetric model and, for the gamma spectrum, on the energy threshold of the telescope. This flux is a bit less than the sensitivity of existing telescopes, so detection is only possible for some models. An even larger amplification ($\approx 10^4$ times Ipser & Sikivie) was suggested by Berezinsky[9], who proposed that the density can vary like $\rho(r) \sim r^{-1.8}$ down to 0.1 pc from the galactic centre. The signature will be a very bright point-like source (angular size smaller than detector resolution) with an expected flux well above the ACT sensitivity.

Extragalactic diffuse light as a probe of the era of galaxy formation

Dark matter is a necessary ingredient to explain structure formation in the early universe, but the scenario is different according to the nature of the dark matter. With hot dark matter large structures appear first and galaxies (and stars) are formed afterwards. In contrast, cold dark matter would trigger galaxy formation earlier. The light emitted by the stars contributes to the infrared/visible extragalactic background light (EBL). The measurement of the density of the EBL can then put constraints on the nature of the dark matter. Unfortunately, direct measurement is very difficult in this wavelength region due to the very dominant foregrounds (zodiacal light, etc...) and only upper limits have been obtained so far.

By using the absorption of gamma rays coming from distant AGN it may be possible to infer the EBL density. The absorption process is electron-positron pair production ($\gamma_{AGN}\,\gamma_{diff} \rightarrow e^+ e^-$) and will appear as a break in the energy spectrum of the AGN. Calculations show that the EBL as a steeply falling spectrum as a function of the energy so that the extragalactic universe is transparent below 10 GeV and completely opaque above 100 TeV. In between, sources at $z = 0.03$ should show a cut-off at a few TeV (scattering on the near infrared) while for sources at $z = 1$ the cut-off is below 100 GeV

(scattering on visible light). Because of the dilation of the universe, the absorption occurs mainly close to the source, so the EBL can be measured as a function of the red shift, z, and constrain evolution scenarios[10].

The fact that only two AGN have been detected in the TeV range, and that these are the among the closest (z = 0.03) objects emitting ganuna rays suggests that this absorption process has already been observed[11]. The EGRET satellite has detected 50 AGN with redshifts ranging from 0.03 to over 2 none of which show any cut-off in the energy spectrum up to 10 GeV[12]. The extrapolated fluxes for these sources at 100 GeV are above 10^{-10} cm^{-2} s^{-1} (for a typical spectral index of 2), thus above the ACT sensitivity. Some projects aim to lower the energy threshold of the ACT down to 20 GeV or less. They should be able to detect many AGNs at different z and answer the question of whether the absorption of gamma rays occurs inside the source (study of the AGN acceleration process) or on their way to Earth (measurement of the EBL).

2. The Atmospheric Cerenkov Technique (ACT)

I will briefly review the main features of ACT telescopes, limiting myself to the imaging and the sampling techniques. I will describe the different strategies to improve the sensitivity and extend the energy range. I will also stress the intrinsic limitation of the technique. The properties of the major experiments in operation and of projected experiments can be found in table I.

Sensitive area

The atmosphere with its 27 radiation lengths is a good electromagnetic calorimeter. A gamma shower starts at 25 km in the stratosphere with a maximum depending on the energy of the primary (around 10 km at 100 GeV). The mean Cerenkov angle is about 1° so that the Cerenkov light is spread over a disk of 250 m diameter for a shower at zenith and for an observatory at an altitude of 1500 m. A telescope anywhere in that disk can detect the shower, so the sensitive area of this detector is around 5×10^4 m^2. For simple geometrical reasons, the size of the Cerenkov pool increases with the distance between the observer and the shower maximum, that is, for showers observed at large zenith angles or for observers at sea level. On the other hand, the Cerenkov light density decreases accordingly, leading to a higher energy threshold. At an elevation of 45° the sensitive area is multiplied by 2 and at 35° by 5. This effect has already been tested by two different telescopes [13, 14].

Energy Threshold and field of view

The energy threshold of a detector is generally dictated by its ability the separate the Cerenkov light from the diffuse background light coming from the sky. Because the Cerenkov light is emitted mostly in the blue band and because the pulse duration lies in the nanosecond range, only fast photo multipliers (PM) can be used. The price to pay is a collection efficiency (ε) reduced to less than 20%. The amount of Cerenkov light on the ground is proportional to the shower energy and the collection area of the telescope (A). Simulations show that for energies around 100 GeV, the Cerenkov density is 0.06 / m^2 / GeV. The density of background light is proportional to the field of view (Ω) and to the integration time (τ) of the PM and the trigger electronics. A shower is detectable if the Cerenkov pulse

Running experiments	Technique	Location	Altitude (km)	Latitude	Collection area (m2)	Pixel size	Field of view	Energy Threshold	Detected sources
Whipple	Imager	Arizona	2.3	31°N	75 + 66	0.25°	3°	250 GeV	Crab, Mrk421, Mrk501
Cangaroo	Imager	Australia	0.2	31°S	11	0.15°	3°	1 TeV	PSR1706-44
Bigrat	Imager	Australia	0.2	31°S	3 x 4	0.25°& 1.6°	2.5°	600 GeV	
Airobicc	Sampling	Canaries	2.2	29°N	49 x 0.13		70°	20 TeV	
C-Hegra	Imager	Canaries	2.2	29°N	5 + 8.5	0.25°&0.5°	3.3°	1 TeV	Crab, Mrk421
Shalon-Alatoo	Imager	Kasahkstan	3.3	42°N	2 x 10	0.7°	8°	> 1 TeV	Crab, Mrk421 (?)
Asgat	Sampling	France	1.6	42°N	7 x 37	1.9°	5°	600 GeV	Crab
Themistocle	Sampling	France	1.6	42°N	18 x 0.5		2.3°	3 TeV	Crab
GT-48	Imaging	Crimea	1.1	45°N	2 x 13	0.4°	2.8°	1 TeV	Crab
Under development									
C-Hegra	Imager	Canaries	2.2	29°N	5 x 8.5	0.25°	3.3° & 4.8°	300 GeV	1997
Mark VI	Imager	Australia		31°S	3 x 50	0.25° & 0.5°	3°	< 250 GeV	1995
CAT	Imager	France	1.6	42°N	15	0.12°	4.8°	200 GeV	1996
Projects									
Magic	Imager	Canaries (?)	2.2	29°N	216	0.08°	2°	50 GeV	1998 (?)
Telescope Array	Imager	Australia (?)			25 x 7	0.25°	8°	100 GeV	(?)
Whipple Upgr.	Imager	Arizona	2.3	31°N	2 x 75	0.125°	6°	100 GeV	1999
Cangaroo II	Imager	Australia	0.2	31°S	38	?	?	100 GeV	(?)
Celeste	Sampling	France	1.7	42°N	168 x 54		0.6°	< 20 GeV	1997 (?)
Stacee	Sampling	California	0.6	35°N	1800 x 40		0.6° ?	< 20 GeV	1997 (?)

Table I : Characteristics of the major experiments in operation, experiments under development and projects.

is well above the background fluctuations so that the energy threshold is given by:

$$E_{th} = \sqrt{\Omega \tau / A / \varepsilon}$$

The basic telescope consists of a large mirror with a PM in the focal plane to detect photons in a small field of view around the mirror axis. From the ground, the shower appears as an elongated ellipse of typically about 0.3° long and 0.1° wide. This ellipse is not centred on the shower axis but can be offset as much as 1.5°. This parallax depends on the altitude of the shower maximum and on the impact parameter of the shower. A small field of view (0.5°) improves the energy threshold with a significant loss in the sensitive area, whereas a large field of view (> 3°) gives the telescope the maximum sensitive area. The only way to reconcile large sensitive area and low energy threshold is to pixelize the camera, that is, to replace a single PM by a large number of small PMs. With this technique, called imaging, the solid angle Ω can be reduced to the image size of the shower. The oldest and still the most powerful imaging telescope is the Whipple Obs. In 1988, the first crude imaging camera was replaced by a finer one, decreasing the energy threshold from 700 GeV to 400 GeV and then to 250 GeV with some other improvements (collecting cones). This movement toward small pixel size was pursued by the CANGAROO collaboration and more recently by the CAT collaboration who hope to reach the same energy threshold as the Whipple with less than half the mirror diameter.

For mechanical reasons, an orientable dish much larger than 10 m is costly. To increase the collection area further one must build an array of independent reflectors. The Cerenkov front will hit the different mirrors at different times so that complicated, and costly, trigger and read-out electronics are needed. No convincing results have yet been obtained with arrays of imagers, but tests are still going on. The sampling technique consists on an array of telescopes with one, or a few, PMs on each. THEMISTOCLE[15] and ASGAT[16] are pioneer experiments of the sampling technique. The small number of channels, compared with arrays of imagers, makes very large arrays of more than 5000 m^2 conceivable. Large collection areas are available for free in decommissioned solar plants; and two projects, CELESTE[17] at Themis and STACEE[18] at solar II have been proposed for feasibility tests. In solar plants, each heliostat focuses the Cerenkov light onto a common solar furnace at the top of a tower. To separate the signals reflected from different heliostats, the furnace is replaced by secondary optics having a PM camera in the image plane such that each PM "sees" only one heliostat. The solar plant is then exactly equivalent to a sampling array. For space considerations and in order to lower the energy threshold, the field of view, defined by the size of the secondary optics, is limited to 0.6°. The shower image must be centred for each heliostat's PM so each heliostat must be aimed at the shower maximum. This constrains the shower axis to be well-centred in the heliostat field, limiting the sensitive area to 10^4 m^2. A 20 GeV threshold should be reached with an array of 40 heliostats.

The collection efficiency is severely limited by the PM quantum efficiency. New photo detectors like AsGaP photocathodes or avalanche photo diodes are promising devices that can increase the quantum efficiency by a factor 2 or more. More technical developments are needed to adapt these detectors to the ACT domain, in particular concerning the range of wavelength sensitivity.

Duty cycle and source visibility

The increase of the background light due to the moon is so great that ACT telescopes do not operate when the moon is up. The duty cycle is thus limited to 5 or 10% depending on the weather. The ARTEMIS collaboration has demonstrated that it is possible to operate with the full moon by restricting the wavelength sensitivity to the UV range (solar blind PMs plus filter). Due to the absorption of UV light in the atmosphere, the energy threshold is increased by a factor 4-5 and the sensitive area is limited to ~ 10^4 m2[19]. Although observation with a UV filter and a classical camera has been tested successfully by the Whipple collaboration[20], this operation mode is limited to nights with a faint moon so the increase of the duty cycle is small.

The lowest energy threshold is obtained for observations at zenith angles less than 35°. For a typical ground-based Cerenkov observatory (lat ~ 30-40°) this represents nearly half of the sky. It is possible to observe sources lower on the horizon but at a higher energy threshold; for example the Crab nebulae was detected from the southern hemisphere by CANGAROO (zenith angle > 53°) at 7 TeV whereas the nominal threshold is 1 TeV[13]. Detection in the northern hemisphere of a gamma-ray line from the galactic centre is thus only possible if the intrinsic threshold of the telescope is very low (~ m_χ / 10). A given source is observable in the optimum zone less than 25% of the time. If we take 10% as a mean value, the observation period for a source is between 50 and 100 hours / year.

Angular resolution and energy resolution

For an imaging telescope the shower image is simply given by the amount of light collected by each pixel. The primary direction lies on the main axis of this image at a distance from the centre of the image that depends on the impact parameter. So, in its simplest version, the imaging techniques allow measurement of one co-ordinate of the primary direction with a typical accuracy of 0.1-0.2°, the other co-ordinate being merely constrained by the limited field of view. Two dimensional reconstruction is possible with a stereoscopic system of 2 or more telescopes separated by 50-100 m. However, at least 3 telescopes are necessary to recover the full sensitive area of a single mirror. The first stereoscopic system was built at Whipple in 1992, but due to mechanical problems with the second telescope, no significant results were obtained. In the Canaries, the HEGRA collaboration is installing a system of 5 small imagers, 80 m apart. The first one and a prototype have given some encouraging results but the analysis of the events detected in coincidence is not yet finished. Exploiting the fact that the exact image shape depends slightly on the impact parameter, simulation shows that full reconstruction is possible with a single imager. This demands a high resolution camera[21]. The lateral distribution of Cerenkov light on the ground is not uniform, so energy resolution also depends on the accuracy of the impact parameter measurement. With a limited field of view (~3°) part of the Cerenkov light is lost outside of the camera and the energy measurement is degraded. Energy resolution should be improved from 40% for the existing imagers to 20% with the second generation telescopes[21,22].

Samplers use the Cerenkov pulse arrival times and amplitudes at each station to determine the direction and the energy of the primary. In the TeV range, the Cerenkov wave front is a cone so

timing information is enough to achieve an angular resolution of 0.1-0.2°[15]. At lower energy, the shower longitudinal development is so short that the wave front becomes spherical and timing information can only give the position of the shower maximum. The impact parameter can be inferred from the lateral distribution on the ground. The resolution is expected to be nearly as good above 50 GeV but will be degraded at lower energy by the geomagnetic field[23]. Note that this method is only valid if the lateral extension of the field is at least as large as the Cerenkov pool. Because the amplitude is measured on many different point of the Cerenkov pool, energy resolution better than 10% can be reached[15].

Sensitivity

Sensitivity to γ rays depends on the ability to reject the diffuse background of protons and, to a lesser extent, of electrons. Images of hadronic showers are broader and longer than electromagnetic ones. Using simple shape cuts, a simple imager can obtain rejection factors between 50 and 150. An array of imagers, looking at the same shower from different directions, should have much larger rejection power. The sensitivity will be ultimately limited by the electron background (figure 1). Samplers can perform similar cuts based on the uniformity of the lateral distribution of the Cerenkov light, but they appear to be much less powerful. At low energy (<100 GeV) proton showers generate less Cerenkov light; compared to gamma showers they generate 1/2 as much at 1 TeV and 1/100 at 20 GeV. With this natural cut, almost all the low energy protons are rejected at the trigger level. In this low energy region, local muons, for which only a fraction of the Cerenkov ring is detected, become a major source of background for single imagers. For point-like sources, pointing criteria can be used to reject the remaining protons and the electrons.

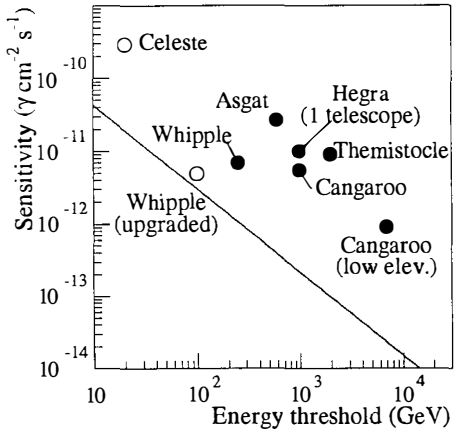


Figure 1 : Sensitivity for point-like sources of the existing (black circles) and proposed (open circles) experiments for an observation time of 50 hours. For extended sources the sensitivity is 3 to 5 times worse. The line shows to the limiting sensitivity (electron background) with a angular resolution of 0.1°.

Conclusions

The nature of the dark matter is still a open question after Zwicky first raised the question in the fifties. The answer might come directly from the observation of new objects —WIMPs, MACHOs,

brown dwarfs, or other— or by the accumulation of indirect information. For both of these approaches, the ACT may play a role.

A direct signature of WHIMPs could be obtained by observing mono-energetic gamma rays from the centre of the Galaxy; this search requires a good ACT detector. Nevertheless, a negative result would not exclude a WIMP hypothesis as long as there exist too large a class of scenarios of dark matter accumulation at a galactic centre.

The probing of the intergalactic diffuse light by energetic gamma rays from distant quasars may help to distinguish between early or late galaxy formation scenarios and, in consequence, between cold or hot dark matter hypothesis. The present (z=0) density of diffuse light is hard to measure because of the foreground light from our Galaxy. Furthermore, the diffuse light variation with z can only be inferred through the absorption of gammas from quasars at different z-values. This study requires that the gamma radiation is observed with a good sensitivity within a large energy domain, typically from 1 GeV to 1 TeV. Such a goal could be fulfilled with simultaneous data from the space detectors EGRET —and/or AMS— and from the ground based detectors CAT and CELESTE.

References

1. Weekes T. et al., Ap. J. 342 (1989) 379. Vacanti G. et al, Ap. J. 377 (1991) 467.
2. Kifune et al., Ap. J. 438 (1995) L91.
3. Punch M. et al., Nature 358 (1992) 477.
4. Catanese M. et al., Padova Workshop on TeV Gamma Rays, September 1995, p.348.
5. Jungman G., Kamionkowski M. & Griest K., Phys. Rep. 267 (1996) 195.
6. Bergström L. & Kaplan J., Astroparticle phys. 2 (1994) 261.
7. Ipser J.R. & Sikivie P., Phys. Rev. D35 (1987) 3695.
8. Urban M. el al., Phys. Lett. B 293 (1992) 149.
9. Berezinsky V., Bottino A. & Mignola G., Phys Lett. B325 (1994) 136.
10. Macminn D. & Primack J.R., Space science review 75 (1994) 430.
11. De Jager O.C., Stecker F.W. & Salamon M.H., Nature 369 (1994) 294.
12. Thomson D. J. et al., Ap. J. S. 101 (1995) 259.
13. Tanimori T. et al., Ap. J. 429 (1994) L61.
14. Krennrich F. et al., Padova Workshop on TeV Gamma Rays, September 1995, p.161.
15. Baillon P. et al. Astroparticle Phys. 1 (1993) 341.
16. Goret P. et al., A&A 270 (1993) 401.
17. CELESTE experimental proposal.
18. Tümer T.O. et al., XXIV ICRC, Rome 1995, Vol. 3, p. 480.
19. X. Sarazin et al., Astroparticle Phys. 4 (1996) 227.
20. Chantell M. et al., XXIV ICRC, Rome 1995, Vol. 2, p. 544.
21. Degrange B., Le Bohec S. & Punch M., NIM (to be published).
22. Sembroski G. et al., XXIV ICRC, Rome 1995,Vol. 3, p.428.
23. Paré E., Space science review 75 (1994) 127.

Résumé

En observant l'Univers par le biais de son émission de gammas à très haute énergie, la technique du Cerenkov Atmosphérique (ACT) peut contribuer à résoudre le problème de la matière noire. Deux possibilités sont étudiées. Les WIMPs peuvent être observés par les gammas générés l'ors de leurs annihilations. En mesurant comment les gammas venant d'AGN sont absorbés par la lumière émise par les galaxies dans le milieu intergalactique, on peut donner des informations sur l'époque de formation de ces galaxies. L'état de l'art de la technique sera exposé ainsi que les développements et projets en France et ailleurs.

Indirect Detection of WIMPs with Cosmic Ray Positrons and Antiprotons: Current Status and Future Prospects

Gregory Tarlé
Department of Physics, University of Michigan, Ann Arbor, MI 48109-1120

Abstract

We review the current status of the indirect detection of WIMPs using cosmic ray antiprotons and positrons. Very recent measurements of the low energy (< 1GeV) antiproton/proton ratio are in excellent agreement with new calculations of secondary production on the interstellar medium. The ratio exceeds the most optimistic values predicted for WIMP annihilation at all energies dimming the prospects for WIMP detection through this channel. Recent measurements of the positron fraction have not confirmed the interesting rise at high energies (> 10 GeV) reported earlier. Greatly improved measurements with higher statistics and improved hadron rejection will be required to uncover even the most optimistic predictions of WIMP annihilation into positron-electron pairs.

Antiprotons

Antiprotons are produced as secondaries through the interaction of cosmic rays (mostly protons) with the interstellar medium. One expects the ratio $\bar{p}/p$ to decrease abruptly below 1 GeV due to kinematic constraints and decline slowly at very high energies due to rigidity dependent leakage from our Galaxy. An early prediction of this ratio calculated within the context of the Leaky Box Model[1] is shown in Fig. 1. The lower bound is the ratio in interstellar space whereas the upper curve includes the effects of solar modulation. Early measurements[2-4] showing an excess of $\bar{p}$s were received with considerable surprise. A number of unusual explanations were proposed to explain the excess. Some models could explain the excess at high energies (closed galaxy[5], continuous acceleration[6], extragalactic sources[7]) or at low energies (evaporation of primordial black holes[8]) but not both. Perhaps the most interesting was the possibility that heavy supersymmetric relic particles might be annihilating in the Galactic halo to produce the excess of low energy antiprotons[9,10]. One particular model[10] shown in Fig. 1 invoked a 15 GeV Higgsino as the lightest supersymmetric particle and was able to reproduce all the early measurements.

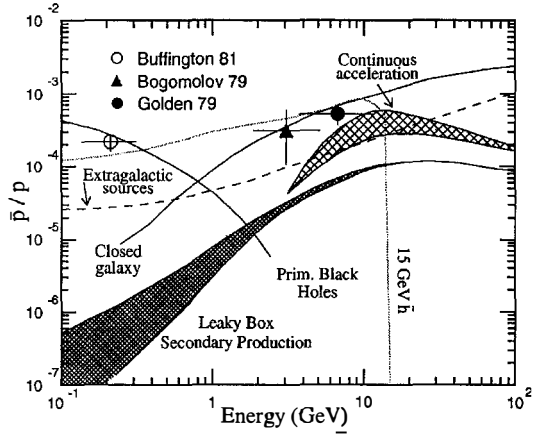


Figure 1. Early measurements of the $\bar{p}/p$ ratio compared to expectations based on secondary production. Also shown are theoretical models created to explain the apparent excess of $\bar{p}$s.

The PBAR experiment[11,12] was the first to employ powerful methods of particle identification using rigidity R vs. time-of-flight (TOF) to produce a mass spectrum. No antiprotons were observed by PBAR and a lower limit was set more than an order of magnitude below ref. 4 (see Fig. 2). LEAP[13] confirmed the PBAR

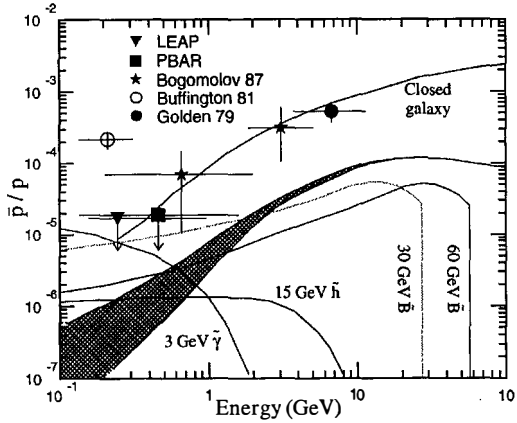


Figure 2. New measurements as of 1987 and the more "realistic" models which add $\bar{p}$s to the secondary flux.

results and Bogomolov et al.[14] continued to see $\bar{p}$s. Theorists began to devise more realistic WIMP models[15,16] which lay tantalizingly below the PBAR measurements.

Within the last year, two new instruments, BESS[17] and IMAX[18], using R vs. TOF but with significantly greater exposure, have identified $\bar{p}$s over the energy range relevant to WIMP detection. In addition, the MASS2 experiment[19], using the same technique as Golden et al.[2], has obtained a much smaller ratio. In addition, new calculations of the expected antiproton and proton flux[20] have been made using improved interstellar reference spectra[21,22] and an improved model of solar modulation. Fig. 3 shows the new measurements, and the new predictions. The dotted lines represent the range of uncertainty using the extrema provided in ref. 21 and the solid line is the most probable calculation using ref. 22 for the interstellar spectra. The calculations are in excellent agreement with the new measurements with the uncertainty in the calculation now exceeding the statistical uncertainty in the measurements. It is evident that the early measurements of antiprotons were dominated by background and that the BESS and

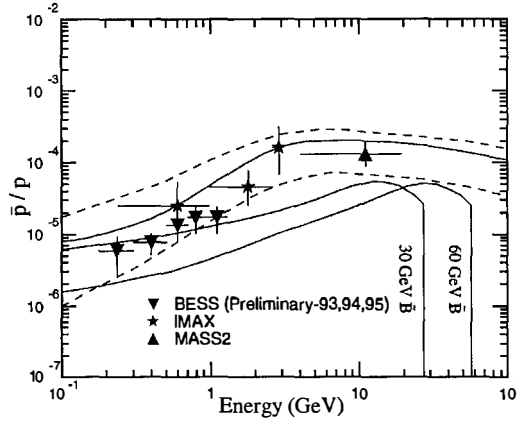


Figure 3. Current $\bar{p}$ measurements with updated secondary models and the WIMP model from ref. 16).

IMAX measurements are the first genuine detection of cosmic ray antiprotons. The current measurements and theoretical expectations make it unlikely that any feature will ever be seen above background. Future measurements of antiprotons will clarify issues of interstellar propagation and the details of solar modulation but will not permit a search for WIMPs.

Positrons

Positrons, like $\bar{p}$s, result from secondary interactions of cosmic rays (mostly protons) with the interstellar medium. At energies below a few GeV the uncertain effects of solar modulation makes interpretation difficult. Between 1-10 GeV the positron fraction $e^+/(e^+ + e^-)$ is low (~10%) indicating a sizable contribution from primary e^-. Above 1 GeV, the positron fraction, is expected to fall slowly with energy[23]. Instead, early measurements[24] (See Fig. 4) showed a dramatic increase in the positron fraction above 10 GeV. A number of theoretical explanations were proposed. Of particular interest was the possibility that WIMPs in the Galactic halo could be annihilating into $e^{\pm}$ pairs, creating the feature in the positron fraction. For WIMPs lighter than the W, the annihilation would proceed directly, producing a distinctive rise followed by a cut-off at the WIMP mass[25]. For WIMPs heavier than the W or Z mass, annihilation would proceed to $W^{\pm}$ or Z^0 pairs[26] which could then decay to produce $e^{\pm}$. This would produce a characteristic "bump" at high energies. Although the most optimistic scenarios did not provide a sufficient $e^{\pm}$ yield to explain the data, the authors of ref. 25 artificially "boosted" the

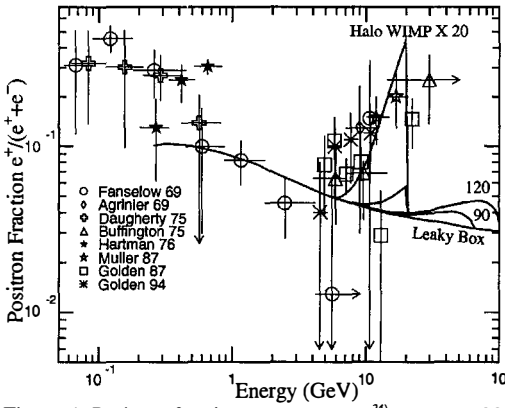


Figure 4. Positron fraction measurements[24] prior to 1995 with predictions from the Leaky Box and various WIMP annihilation models.

calculations by a factor of 20 as shown in Fig. 4 to "make comparison with the data intriguing."

Since 1995, three new positron measurements (HEAT[27], MASS2[28]and TS93[29]) have been made using advanced methods of particle ID to eliminate hadron background.. The HEAT experiment[27] utilized a magnetic spectrometer, transition radiation detectors, an electromagnetic calorimeter and energy momentum matching to achieve an unprecidented level of hadron rejection at high energies. The HEAT data shown in Fig 5 do not show the dramatic rise seen previously. It is apparent that earlier measurements were dominated by hadronic background.

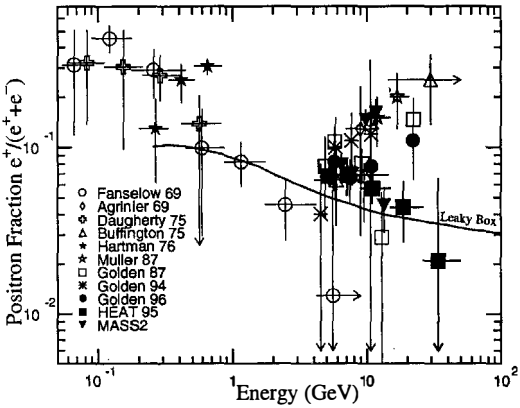


Figure 5. Current positron fraction results. The HEAT data are consistent with the simple Leaky Box model.

Detection of WIMPs through the positron channel will present a unique experimental challenge in the future. Although the detection of a "bump" at high energies would represent a "smoking gun" for WIMP annihilation, the models could easily accommodate several orders of magnitude reduction in the signal. Even in the most optimistic scenario shown in Fig. 4, larger exposures and improved particle ID are required. It is possible that long duration balloons and/or dedicated satellites such as PAMELA[30] may make this possible within the next five years.

References

1) Protheroe, R.J., Ap.J. **251**, 387 (1981).
2) Golden, R.L., *et al.*, Phys. Rev. Lett. **43**, 1196 (1979).
3) Bogomolov, E.A. *et al.*, Proc.16th Inter. Cosmic Ray Conf. (Kyoto) **1**, 330 (1979).
4) Buffington, A. *et al.*, Ap. J. **248**, 1179 (1981).
5) Stephens, S.A., Ap. Space Sci. 76, 87 (1981).
6) Simon, M., *et al.*, 20th Inter. Cosmic Ray Conf. (Moscow) **1**, 330 (1979).
7) Stecker, F.W., Protheroe, R.J., and Kazanas, D., Ap. Sp. Sci. **96**, 171 (1983).
8) Turner, M.S., Nature **297**, 379 (1982).
9) Silk, J.and Srednicki, M., Phys. Rev. Lett., 53, 624 (1984)

10) Hagelin, J.S., Kane, G.L. and Raby, S., Nucl. Phys **B241**, 638 (1984).
11) Ahlen, S.P. *et al.*, Phys. Rev. Lett. **61**, 145 (1988).
12) Salamon, M.H. *et al.*, Ap. J. **349**, 78 (1990).
13) Streitmatter, R.E. *et al.*, Adv. Sp. Res. **9**, 65 (1989).
14) Bogomolov, E.A. *et al.*, Proc.20th Inter. Cosmic Ray Conf. (Moscow) **2**, 72 (1987).
15) Ellis, J. *et al.*, Phys. Lett. **B214**, 405 (1988).
16) Jungman, G. and Kamionkowski, M., Phys. Rev. **D49**, 2316 (1994).
17) Yoshimura, K. *et al.*, Phys. Rev. Lett. 75, 3792 (1995) and J. Ormes, private communications.
18) Mitchell, J.W. *et al.*, Phys. Rev. Lett. 76, 3057 (1996).
19) Hof.,M. *et al.*, Proc.24th Inter. Cosmic Ray Conf. (Rome) **3**, 60 (1995).
20) Labrador, A.W. and Mewaldt,R.A., Proc.24th Inter. Cosmic Ray Conf. (Rome) **3**, 68 (1995).
21) Gaisser, T.K. and Schaeffer, R.K., Ap. J. **394**, 174 (1992).
22) Webber, W.R. and Potgeiter, M.S. Ap. J. **344**, 779 (1989).
23) Protheroe, R.J., Ap. J. **254**, 391 (1982).
24) Fanselow J.L. *et al.*, Ap. J. **158**, 771 (1969); Agrinier, B. *et al.*, Lett. Nuovo Cimento **1**, 153 (1969); Daugherty, J.K. *et al.*, Ap. J. **198**, 493 (1975); Buffington, A. *et al.*, Ap. J. **199**, 669 (1975); Hartman, R.C. and Pellerin, C.J., Ap. J. **204**, 927 (1976), Müller, D. and Tang, K., Ap. J. **312**, 183 (1987); Golden, R.L. *et al.*, Astron. Astrophys. **188**, 145 (1987), Golden, R.L. *et al.*, Ap. J.. **436**, 769 (1994).
25) Turner, M.S. and Wilczek, F, Phys Rev. **D42**, 1001 (1990).
26) Kamionkowski, M. and Turner, M.S., Phys Rev. **D43**, 1774 (1991).
27) Barwick, S.W. *et al.*, Phys. Rev. Lett. **75**, 390 (1995).
28) Basini, G. *et al.*, Proc.24th Inter. Cosmic Ray Conf. (Rome) **3**, 1 (1995).
29) Golden, R.L. *et al.*, Ap. J.. **457**, L103 (1996).
30) Addriani, O. *et al.*, Proc.24th Inter. Cosmic Ray Conf. (Rome) **3**, 591 (1995).

Détection Indirecte des WIMPs par les Positrons et Antiprotons Cosmiques: Situation Actuelle et Perspectives d'Avenir

Abstract

Nous passons en revue la situation actuelle concernant la détection indirecte des WIMPs utilisant les antiprotons et positrons présents dans la radiation cosmique. De très récentes mesures du rapport antiproton/proton à basses énergies (< 1 GeV) sont en excellent accord avec de nouveaux calculs de production secondaire sur le milieu interstellaire. Le rapport dépasse les valeurs les plus optimistes prédites pour l'annihilation des WIMPs à toutes énergies, réduisant la probabilité de détection des WIMPs par cette méthode. De récentes mesures de la fraction de positrons n'ont pas confirmé l'augmentation intéressante précédemment rapportée à hautes énergies (> 10 GeV). Il faudra des mesures grandement améliorées, avec de plus amples échantillons statistiques et un meilleur rejet des hadrons, avant d'atteindre les prédictions les plus optimistes pour l'annihilation des WIMPs en paires d'électron et positron.

"HERON" AS A DARK MATTER DETECTOR?

J. S. Adams, S. R. Bandler[1], S. M. Brouer[2], C. Enss[1], R. E. Lanou, H. J. Maris,
T. More and G. M. Seidel
Department of Physics, Brown University, Providence, RI 02912, USA

Abstract

"**HERON**", which is the acronym for "**He**lium: **Ro**ton detection of **N**eutrinos", is a project whose principal goal is a next generation detector of solar neutrinos from the p-p and ^{7}Be branches. It will utilize superfluid helium as the target material and employ event energy transport out of the target by phonon and roton processes unique to helium. Many of the challenges presented for dark matter detection are very similar to those for low energy solar neutrinos. We present new results from our feasibility studies for **HERON** which indicate an asymmetry in the roton emission distribution from stopping particles and the ability to detect simultaneously the ultra violet fluorescence photons also emitted. These features are potentially valuable for solar neutrino detection and the question is explored as to whether or not the same helium technique could be valuable for WIMP dark matter detection.

Supported in part by D.o.E. DE-FG02-88ER40452
and N.S.F. - 9420744.

[1]Present address: Universitat Heidelberg, Germany.
[2]Present address: Universitat Karlsruhe, Germany.
Presented by: R. E. Lanou (email: lanou@physics.brown.edu)

Measurements of the recoil energy spectrum from low energy solar neutrinos place many of the same demands on any detector, designed to do so, as they do in the case of measurements for WIMP dark matter. Both require an effective mix of several features. Among them are: a) an extremely low radioactivity in the target material and its container, b) good event signature discrimination, c) modest energy resolution, d)real time detection and e) relatively massive targets. Many avenues for event signature discrimination can be valuable; such as event position in the detector by taking advantage of differing mean free paths for interaction by background entering from the exterior, directionality of the recoil to take advantage of sidereal effects or source location, and particle identification to distinguish between nuclear and electron recoils.

We have been developing a particle detection technique based upon the use of superfluid helium as the target material for detecting neutrinos from the p-p and ^{7}Be branches of the sun's principal fusion cycle. The project is referred to as **HERON** (for **He**lium: **Ro**ton detection of **N**eutrinos) which, if the present R&D shows it to be feasible, is intended to be a "next generation" solar neutrino detector detecting > 20 events per day in real time. Such a full scale detector would utilize roughly 10 tons (fiducial volume) of liquid helium in the 100 milli-Kelvin temperature range and the events would consist of elastic scattering of neutrinos by electrons in which the recoiling electron energies are in a continuum up to 660 keV. Liquid helium is a very attractive material for this type of application for many reasons. It is the purest material known since nothing is soluble in it, foreign elements freeze out on the walls (it is self-cleaning), its first nuclear excited state is ~ 20 MeV and it has no long-lived isotopes. It is very inexpensive even in these quantities and industrial/commercial handling on this scale is standard and routine. Additionally, some of the materials suitable for containment cryostats, such as OFHC copper, tend to have few long-lived cosmogenically induced activities. In spite of this attractiveness, the traditional methods for extracting a signal from a fluid target such as drifting ions, reconstruction from scintillation, or calorimetry do not prove adequate to obtaining useful information. The primary problems, respectively, are due to too low drift velocity, too long lifetime for the fluorescing state and too large a heat capacity. However, there are a sequence of phonon and roton processes in superfluid helium which, when taken together, have been suggested [1)] as a means for extracting the energy and other information from a large detector[2)] of just this type. In recent years we have been carrying

out experiments [3] which have established the basic properties of particle detection by this method.

These experiments have been carried out in prototype cells (typically containing 3 liters of liquid helium) attached to a ^{3}He - ^{4}He dilution refrigerator. The interior of the cell contains an array of instrumentation which can be readily re-configured to suit the goals of a particular test. Radioactive sources of alpha particles (< 6 MeV), gamma rays (662 keV) and electrons (364 keV) as well as pulsed heaters are utilized to study the sequence of processes which permit particle detection via phonons. Small superconducting motors are used to move the sources within the liquid which is typically at 30 mK.

While the principal interest for possible application of this technique for WIMP dark matter centers upon our recent experimental results which suggest a sensitivity to recoil track spatial orientation[4] and simultaneous detection of a fast fluorescence signal calorimetrically[5], it is instructive to first describe the sequence of processes which are the basis of the detection method. The following steps occur.

a) A recoil particle > 100 eV generates mainly secondary ions and electrons.

b) The secondary electrons lose energy by further ionization or atomic excitation until they fall below ~20 eV whereupon they scatter from the atom as a whole.

c) recombination of the ionization then occurs resulting in the formation of dimers which subsequently fluoresce primarily ~ 16 eV photons to which helium is transparent.

d) The energy received by atoms scattered from secondary electron collisions appears as phonons and rotons.

e) The rotons (which are a class of phonons distinguished by their position in the non-linear, higher energy portion of the dispersion curve for superfluid helium) dominate the available phase space, are stable against interaction or decay and propagate at ~ 150 m-s^{-1}.

f) The energy of the rotons is greater than the binding energy of a helium atom to the liquid (0.65 meV) and those satisfying certain kinematic conditions undergo a process referred to as quantum evaporation in which a single roton ejects a single helium atom.

g)A consequence of this kinematic condition is that the ejected atoms arise from those rotons contained within a cone (with axis normal to the free surface of the liquid). In our experiments the cone half-angle is ~17^0 and corresponds to about 1% of the recoil energy.

h)The number of ejected atoms is very large (~ 10^5/keV of recoil energy) and these are positioned just above the liquid. The binding energy to the low heat capacity wafers is

~ 9 meV/atom thus producing an energy deposit in the wafer roughly ten times that of the original roton or, for the complete deposition, ~10% of the recoil particle's energy. Additionally, the ultra violet radiation (5 photons/keV) is absorbed in the wafer as a fast pulse.

i) In the final step, tiny thermometers (thin, superconducting films or thermistors) attached to the wafers record the resulting temperature pulse induced.

In any large scale application of this technique a close-spaced, regular array of wafers would be placed just above the liquid and, from the wafer hit pattern as well as pulse magnitude and timing distributions, the coordinates and energy of the event would be derived.

In order to investigate further the possibility of obtaining directional information on the recoil as well as to make measurements of the presence of detectable fluorescence photons, we have performed the following test. A collimated source of alpha particles was constructed; the average energy of the emerging alpha particles was 3.3 MeV with a 15% FWHM determined by the geometry and Kapton window thickness. The source (4.5 cm. below the surface) was mounted on the drive shaft of a small superconducting stepper motor oriented so that the direction of the alpha particles could be rotated in a plane perpendicular to the helium liquid surface (see Figure 1). A 1cm. by 2 cm. silicon wafer with an Ir-Au superconducting transition edge thermometer attached was placed just above the liquid and readout by a SQUID. Data were taken as the track direction was stepped at 6.5^0 intervals through a full 180^0. In Figure 2 are shown traces (100 event averaged to illustrate detail) taken with the track orientation a) horizontal and b) vertical. The rise time of the wafer is ~50 μ-s and the relaxation time to the reservoir is ~500 μ-s. The fast, small step at the start is due to the arrival of the UV photons and the large subsequent rise is due to the slower rotons and is consistent with their known speed of propagation. The value of the pulse heights for both portions of the pulse at all of the track angles is shown in Figure 3. As can be seen, the roton pulses range a factor ~3.5 between a direction parallel and normal to the surface while the photon pulses are independent of track direction as would be expected for photon emission. In contrast, the large difference in energy collection due to rotons suggests a commensurate asymmetry in the radiation of rotons from the track itself. This asymmetry is believed to arise due to the very high density of rotons in the volume swept out by the stopping track and by the large roton-

roton scattering cross section (10^{-14} cm^2). For these alpha particles there are ~10^9 rotons in ~10^{-13} cm^3 resulting in the down conversion of the original rotons to lower energy and then radiating the resulting thermal distribution whose directional intensity reflects the geometrical aspect ratio of the track volume. We have confirmed this thermalization effect in separate experiments previously reported[3]. This picture is also supported by our detection of the fluorescence photons. Experiments[5] have established that the fluorescence results due to collisional de-excitation of helium dimers which requires a region of high energy density. Additionally, the photon intensity we measure (5 photons/keV) is consistent with that measured in these same experiments.

How could either of these effects be used in a real detector of WIMPS or solar neutrinos? Differences in the ratio of scintillation light relative to roton signal strength for electron or ^{4}He recoils may provide a useful particle I.D. Should these observations of roton emission asymmetry survive to lower energy recoils and to other particle types (e.g., electrons) then,with the known distribution of emission versus track orientation and the differential pulse height distribution on the wafers, both the track orientation and the event energy could be found. This would be a very powerful addition to event signature discrimination by correlation of the sun's position with solar neutrino events or to observe the sidereal direction change of WIMP dark matter due to the Earth's motion through the dark matter distribution.

Much work remains before the feasibility of these applications is known. We are now beginning a new set of experiments which will attempt to address some of the issues involved. Among them are experiments on the following. We need to improve the sensitivity of our wafer calorimeters and their attached thermometers in order to build detectors large enough for solar neutrinos and at the same time bring our knowledge of the energy deposition by low energy electrons up to the same level we have for alfa particles (particularly with respect to directionality). In order to get a measure of thepotential for energy, position and directional resolution we are planning to construct a new prototype roughly ten times larger than our present one. It would be 30 liters (~4 kg.) and be instrumented with eight wafer channels. The prototype will also be useful for other

test of a more structural sort such as the use of cesium films for superfluid film flow control and test of construction materials. On this scale, it is not inconceivable that the prototype could itself serve a useful purpose for dark matter detection.

On any scale, however, we must take into account what the weaknesses and strengths of helium are as a target material for WIMP dark matter. Among its weaknesses must be counted its poor cross section for supersymmetric particles (a well motivated candidate for WIMP dark matter); further it is a spin-0 nucleus. Its low nuclear mass is both a weakness and a strength in that although that makes the visible recoil energy greater at the same time it is most effective when the dark matter mass is also low. On the strength side, we can count its extreme purity, the potential for event directionality and position, the absence of form factors at low collision energies and the ease of making very large targets. We will have to wait and see whether our future tests and Mother Nature together can answer our original question.

(See text for figure discussion)

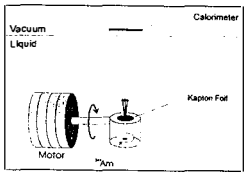


Figure 1

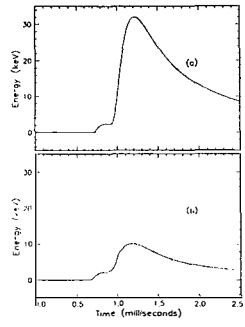


Figure 2

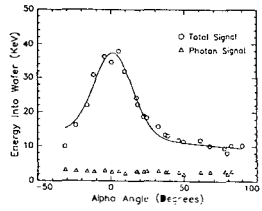


Figure 3

1) R. E. Lanou, H. J. Maris and G. M. Seidel, Phys. Rev. Lett. **58** (1987) 2498.
2) S. R. Bandler, C. Enss, G. Goldhaber, R. E. Lanou, H. J. Maris, T. More, F. S. Porter and G. M. Seidel, J. Low Temp. Phys. **93** (1993) 785.
3) S. R. Bandler, R. E. Lanou, H. J. Maris, T. More, F. S. Porter, G. M. Seidel and R. H. Torii, Phys. Rev. Lett. **68** (1992) 2429; S. R. Bandler, C. Enss, G. Goldhaber, R. E. Lanou, H. J. Maris, T. More, F. S. Porter and G. M. Seidel, J. Low Temp. Phys. **93** (1993) 715.
4) We are indebted to Dr. Susan Cooper of the Max Planck Institute for Physics (Munich) and Prof. F. v. Feilitzsch of the Technical University (Munich) for providing us with this device.
5) M. Stockton, J. W. Keto, W. A. Fitzsimmons, Phys. Rev. Lett. **24** (1970) 654; Phys. Rev. A 5 (19972) 372; and the references therein.

DARK MATTER IN EARLY-TYPE GALAXIES

W. Forman, L. David, and C. Jones
Harvard-Smithsonian Center for Astrophysics
60 Garden St.
Cambridge, MA 02138 USA

ABSTRACT: We describe recent progress in understanding the dark matter content of early-type (elliptical and S0) galaxies. ROSAT observations of the hot X-ray coronae around early-type galaxies show that these galaxies include a dark matter halo with a mass-to-light ratio of $\sim 100 M_\odot/L_\odot$. We compare these measurements to those for larger systems, groups and clusters. We summarize the cosmological implications of the dark matter, hot gas, and stellar mass distributions.

I. Introduction

The history of the study of gas in early type (elliptical and S0) galaxies began at least as early as 1957. In that year Sandage (1957) estimated that during the course of normal stellar evolution, the constituent stars of a bright elliptical would shed $5 \times 10^9 M_\odot$ of gas. Sandage suggested that much of this gas could be in the form of neutral hydrogen although a portion of it could be ionized and might explain the optical emission lines of forbidden oxygen which were seen in 15% of the ellipticals observed by Humason, Mayall, and Sandage (1956).

In the 30 intervening years, the limits on various forms of gas in early type galaxies became

increasingly more stringent and models were developed to explain the absence of gas in elliptical galaxies. For example, Mathews and Baker (1971) discussed the possibility of driving the gas from galaxies in galactic winds powered by supernova explosions.

With the launch of the Einstein Observatory and the ability to obtain spatially resolved X-ray images on galaxy scales, it became apparent that in fact E and S0 galaxies were not devoid of gas, but that the gas was stored in a reservoir of hot, $\sim$1 keV ($\sim 10^7$K) gas (Forman, Jones, and Tucker 1985; see also Nulsen et al. 1984 and Canizares et al. 1987). The quantity of gas was roughly that calculated by Sandage, 25 years earlier.

The gas can tell us something fundamental about early type galaxies — namely their gravitational mass and the shape of the underlying potential which have been difficult to determine for this class of galaxy. The presence of a cool interstellar medium has been used to systematically measure the dynamical masses in spiral systems at large radii with both optical and radio techniques. The absence of such tracers at large radii in early-type galaxies has made the determination of their mass distributions significantly more difficult using optical or radio techniques. However, the observed X-ray coronae provide ideal test particles to trace the total mass distribution in individual early type galaxies. The gas obeys well-understood laws and comprises a small fraction of the total galaxy mass.

The method used to derive the mass distribution in early type galaxies relies on two simple principles — the ideal gas law and hydrostatic equilibrium. The hydrostatic equation can be written as

$$\frac{dP}{dr} = -G\rho M(< r)/r^2 \tag{1}$$

where P is the pressure, ρ is the gas density, G is the gravitational constant, and $M(< r)$ is the total mass interior to the radius r. We should emphasize that the mass in equation 1) is *not* the mass in gas but all the mass that contributes to the gravitational potential whether it be stars, collapsed objects, or exotic particles. The ideal gas law allows us to eliminate the pressure in favor of the gas temperature and density and solving for the mass interior to r we find:

$$M(< r) = -\frac{kT}{G\mu m_p}\left(\frac{d\ln\rho}{d\ln r} + \frac{d\ln T}{d\ln r}\right) r. \tag{2}$$

where T is the gas temperature at a radius r, μ is the mean molecular weight and m_p is the mass of a proton. Numerically this becomes:

$$M(< r) = 3 \times 10^{12}(T/10^7 K)\left(-\frac{d\ln\rho}{d\ln r} - \frac{d\ln T}{d\ln r}\right)(r/100kpc)M_\bullet. \tag{3}$$

If one can measure the temperature and density gradients, then one can measure the total mass

as a function of radius. This is a remarkably powerful technique and has been applied to a variety of systems which are sufficiently massive to bind gas which radiates at X-ray energies.

II. Masses of Early-Type Galaxies

The first X-ray mass determination of an early-type galaxy was performed using Einstein observations of the galaxy NGC4472 Figure 1 shows a ROSAT X-ray contour map superposed on an optical image. The X-ray emission extends well beyond the optical image of the galaxy. To measure the mass of NGC4472 requires the determination of the gas density and temperature profiles. The gas temperature profile as measured from ROSAT PSPC images is shown in Figure 2. The earlier Einstein temperature determinations are also presented and the figure shows the dramatic improvement in precision available with ROSAT compared to the first Einstein observations. We see an isothermal corona with a temperature of about 1 keV (1.2×10^7K) measured to 16 arcminutes (74 kpc) from the galaxy center. Note that there is a small, but significant temperature decrease at small radii where the gas density is sufficiently high and the cooling time is sufficiently short (see Thomas 1986 for a discussion of cooling cores in hot coronae).

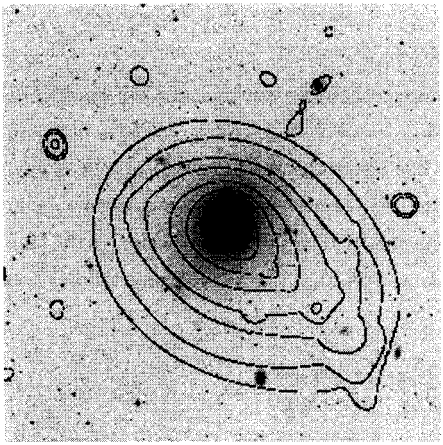


Figure 1 shows an optical image of NGC4472 with the ROSAT PSPC X-ray contour map superposed. The X-ray contours extend well beyond the optically luminous region of the galaxy. The contour map was generated with a new wavelet smoothing technique (Vikhlinin et al. 1996).

The surface brightness can be used to derive the gas density profile which is also needed to calculate the gravitating mass. Figure 3 shows the surface brightness profile. The surface brightness can be written as $S(r) = S_0 \int n^2 \Lambda(T) dV$ where $\Lambda(T)$ is the emissivity of the gas as a function of temperature and n is the radial gas density profile (assuming spherical symmetry). As Figure 2 shows, the gas is essentially isothermal and hence $\Lambda(T)$ is a constant. Fitting the points beyond 2′ to a power law gives the slope of the surface brightness profile, α, which is directly converted to the logarithmic derivative of the gas density, γ, needed in equation 1): $\gamma = (\alpha + 1)/2 = 1.65 \pm 0.15$.

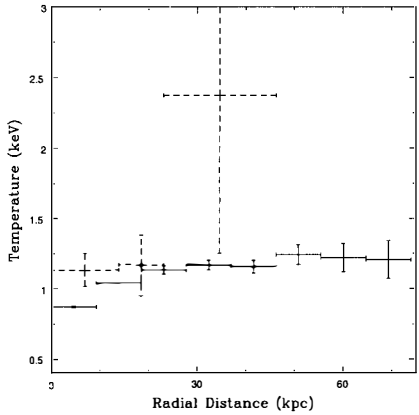


Figure 2 – The radial temperature profile of NGC4472 is shown as derived from PSPC ROSAT observations (solid lines). The temperature decrease at the center may arise from gas with a short cooling time. Also shown are the Einstein observations (dashed lines) which demonstrated that the early-type galaxies had massive, dark halos.

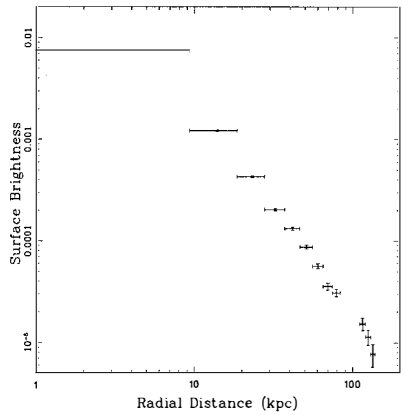


Figure 3 – The radial X-ray surface brightness profile of NGC4472 is shown extending to large radii. Given the isothermal temperature profile, the surface brightness, beyond the central cooling region, translates directly into the gas density distribution which is needed to derive the gravitating mass.

With the measured gas density and temperature profiles, we can calculate the gravitating mass using equation 2. For NGC4472, we use $T_{gas} = 1.3 \times 10^7$K and $\gamma = 0.55 \pm 0.05$ and find a gravitating mass of $4.5 \times 10^{12}\mathrm{M}_\odot$ at 70 kpc where the gas temperature is measured or $6.4 \times 10^{12}\mathrm{M}_\odot$ at 100kpc if we assume that the gas temperature remains constant within the regions where the brightness profile is well-determined. The galaxy luminosity (in the blue band) is $L_B = 7.4 \times 10^{10}\mathrm{L}_\odot$for the assumed distance of 16 Mpc. Taking the uncertainties into account this yields a mass to light ratio of

$$M/L = \begin{cases} 78 - 93 M_\odot/L_\odot & \text{at 100 kpc} \\ 54 - 65 M_\odot/L_\odot & \text{at 70 kpc} \end{cases} \quad (4)$$

These numbers are to be compared to the mass-to-light ratio where the mass is dominated by stellar matter. For example, Lauer (1985) studied a sample of 14 elliptical galaxies. The derived mass-to-light ratios for the central regions of the galaxy are:

$$M/L = \begin{cases} 14 M_\odot/L_\odot & \text{galaxy core} \\ 11.9 M_\odot/L_\odot & \text{within 6.1 kpc} \end{cases} \quad (5)$$

Thus, we clearly see that the outer regions of the galaxy are dominated by dark matter. If we exclude the inner 6.1 kpc, and compute the mass-to-light in the outer region (excluding the

central region of the galaxy) 16-70 kpc, we find a mass-to-light ratio of 96-118 $M_\odot/L_\odot$. Thus, the outer regions are dominated by dark matter and have a very different mass-to-light ration than the central regions.

NGC4472 is not an atypical galaxy. The bright galaxies that have been studied all yield similar results. For example, from work in progress, we find mass-to-light ratios in solar units of 97, 89, 110, and 50-90 (including diffuse light seen at large radii around the galaxy) for NGC507, NGC499, NGC5846, and NGC1399 respectively. For NGC4636, $M/L_B = 145$ at about 100 kpc (Mushotzky et al. 1996).

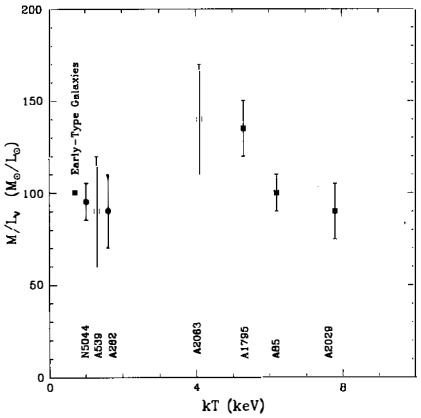


Figure 4 – The mass-to-light ratios for systems ranging from early-type galaxies (within about 100kpc) to rich clusters (measured within 1 Mpc). The error bars include only the uncertainty in the total mass as derived from the X-ray observations. The data show that the mass-to-light ratios are nearly constant over a wide range of systems and demonstrate that the dark matter in clusters is very likely the same material in galaxy halos.

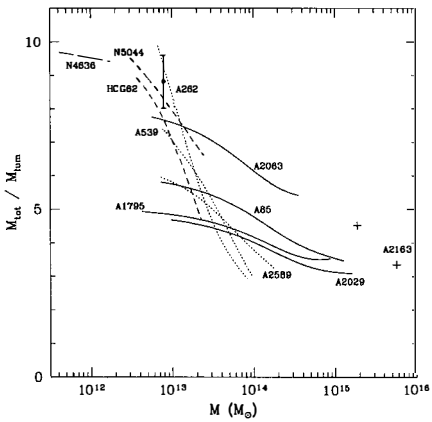


Figure 5 – The ratio of total gravitating mass to luminous mass (including both hot gas and stellar matter) as a function of total gravitating mass. A typical elliptical, groups, cool clusters, and hot clusters are shown long dashed, short dashed, dotted, and solid lines respectively. The two symbols (+) show measurements for A2163 made at two different radii. On the largest mass scales, the ratio of total to luminous baryonic matter is about 3-4.

III. Comparison of Galaxies to Groups and Clusters

How do the galaxies compare to larger/hotter systems. Figure 4 shows the mass-to-light ratios for a variety of systems from galaxies to rich clusters. The figure shows a remarkable constancy of the ratio of dark to luminous matter on scales from galaxies to rich clusters. The linear scales range from 100kpc for the galaxies to 1 Mpc for the clusters. This suggests that the dark matter in clusters is the same as that in galaxies. The dark matter in clusters may be only the dark matter stripped from galaxy halos. No new dark matter component is required in clusters.

With the ability to determine the gas, stellar, and gravitating masses, we can compare the observable (luminous) baryonic matter to the total gravitating matter. In galaxies, the gas

represents only a few percent of the total mass and the stellar light dominates the observed baryonic matter. In rich clusters, the hot intracluster gas is the dominant baryonic component (3-5 times more massive than the stellar matter observed in galaxies). Figure 5 shows the variation of the ratio of total mass to luminous mass (the inverse luminous baryonic fraction). From a high value of 10 for galaxies, the value decreases by more than a factor of two. Thus, as one goes to larger scales, the universe is brighter (as long as one takes a panchromatic view).

Finally, we mention that Figure 5 suggests that the gas mass fraction (equivalently the minimum baryonic fraction) on cluster scales approaches $f_b = \Omega_b/\Omega \sim 0.30 h_{50}^{-3/2}$. If this fraction is representative of the universe as a whole, then, with the Big Bang Nucleosynthesis constraint that $0.04 < \Omega_b h_{50}^2 < 0.06$ (Walker et al. 1991), we can derive an upper limit on the total mass density of $\Omega < 0.2 - 0.3$ (see White et al. 1992 and David et al. 1995). This result argues in favor of a low-density universe.

IV. Conclusions

The application of X-ray imaging to the study of early-type galaxies and the discovery of their hot gaseous corona has provided a new tool for measuring the distribution of dark matter around these galaxies. The X-ray observations show that early-type (elliptical and S0) galaxies, like their spiral counterparts, are surrounded by dark halos with mass-to-light ratios of ~ 100. The luminous baryonic fraction increases from about 10% in galaxies to a roughly universal value of about 30% of the total mass for rich clusters. The large baryonic fraction observed in hot gas and the limit on the baryonic fraction from Big Bang Nucleosynthesis argue for a low density universe.

V. References

Canizares, C. R., Fabbiano, G. and Trinchieri, G. 1987, ApJ, 312, 503

David, L., Jones, C. and Forman, W. 1995, ApJ, 445, 578

Forman, W., Jones, C., and Tucker, W. 1985, ApJ, 293, 102

Humason, M.L., Mayall, N.U., and Sandage, A.R. 1956, AJ, 61, 97

Lauer T. 1985, ApJ, 292, 104

Mathews, W. and Baker, J. 1971, ApJ, 170, 241

Mushotzky, R. et al. 1996, preprint

Nulsen, P.E.J., Stewart, G., and Fabian, A.C. 1984, MNRAS, 208, 185

Sandage, A. 1957, ApJ, 125, 422

Thomas, P. A., Fabian, A. C., Arnaud, K. A., Forman, W., Jones, C. 1986, MNRAS, 222, 655

Vikhlinin, A., Jones, C. and Forman, W. to be submitted to ApJ

Walker, T., Steigman, G., Schramm, G., Olive, K., and Kang, H. 1991, ApJ, 378, 186

White, S., Navarro, J., Evrard, A., and Frenk, C. 1993, Nature, 366, 429

SUSY SM AND SUSY GUT AXIONS AS DARK MATTER

Jizong Lu [1]
Institut für Physik, Universität Dortmund, D-44221 Dortmund, Germany
and
Department of Physics, Shanghai Teachers University, Shanghai 200234, P.R. China [2]

Abstract

Recently the interest in the axion has enhanced since it is a good candidate for cold dark matter. We discuss invisible axions, which agree with conditions as a cold dark matter candidate, in both local SUSY SU(3)⊗SU(2)⊗U(1) and local SUSY GUT SO(10) models. Several possible ways of searching for axions are discussed too.

[1]DAAD/K.C. Wong Fellow; e-mail address: luj@hal1.physik.uni-dortmund.de
[2]permanent address

1. Introduction

There are several obvious evidences indicating the existence of nonluminous matter in the Universe, galaxies and holes, furthermore, the bulk of it must be non-baryonic.[3] The interest in axions as a possible dark matter candidate has been enhanced recently, although there are many others. There are two reasons. First, the COBE observations of structure in the microwave background radiation favor cold dark matter cosmologies [1]. The axion is one of prime candidates for cold dark matter. Second, some interesting experiments to search for the axion have a realistic chance of finding these elusive particles [2].

The axion was first proposed as an elegant solution to the strong CP violation problem in QCD [3]. It was realised soon that axions could have an important role to play as a dark matter candidate since they acquire a small mass at the QCD phase transition.

In this talk I would like to present our results of axions in both SUSY SM and SUSY GUT models, in which the strong CP violation was eliminated via the DFSZ mechanism [4] and the $U_{PQ}(1)$ breaking scale f_a is close to the geometric hierarchy mass scale $M_R \equiv \sqrt{M_U M_W} \approx 10^{11}\sim 10^{12}$GeV. Thus, the axion mass in the models is around 10^{-5}eV and the low energy coupling to normal matter is suppressed by $1/f_a$. These make it to be a good candidate for cold dark matter. There are only two prime cold dark matter candidates: axions and neutralinos [5], so to discuss axions as cold dark matter in some SUSY models is still significant.

2. The Axion in a SUSY SU(3)⊗SU(2)⊗U(1) Model

There are several motivations to discuss axions in SUSY models. Say, there is so called automaticity of the PQ symmetry. Two Higgs doublets are needed to implement a $U_{PQ}(1)$ symmetry in original axion models. On the other hand, minimal SUSY models also require two SU(2) doublet Higgs superfields H and H'. Because if there were only one Higgs superfield, the theory could not be anomaly free. SUSY might give PQ symmetry an automaticity. Thus, the interest in SUSY axion models has recently soared [6].

At first, I would like to introduce a SUSY $SU_c(3)\otimes SU(2)\otimes U(1)$ axion model, in which the gauge interaction is the same as those in the Standard Model (SM). The superpotential f is

$$f = aQU^cH + bQD^c\bar{H} + cLE^c\bar{H} + dH\bar{H}S + \delta SSS, \tag{1}$$

where H, $\bar{H}$, L and Q are SU(2) doublets; D^c, U^c and E^c are singlets; Q, D^c and U^c are SU(3) triplets and S is the singlet for both SU(3) and SU(2). The VEV's are $< S >= v, < H >= v_1, < \bar{H} >= v_2$ with all others vanishing. We can set $v_1 \sim v_2 \ll v$. The potential ν is

$$\nu = |\frac{\partial f}{\partial z_i}|^2 + |g_\alpha z_i^* T_\alpha z_i + \xi_\alpha|^2, \tag{2}$$

where z_i denote all fields, g_α is the group coupling constant and T_α is the group generator. ξ_α is an arbitrary parameter which is non-zero only for $U_Y(1)$ group and to break the supersymmetry. There is an additional $U_A(1)$ symmetry in the potential ν. Axions will be occur in its breaking.

[3]for example, see lectures: K. Olive, *Why Do We Need Non-Baryonic Dark Matter* and G. Jungman, *Particle Dark Matter* in this proceedings.

If replace the complex scarlar fields $H, \bar{H}$ and S by their phase fields respectively: $H_0 \to e^{iv_1\eta_1}$, $\bar{H}_0 \to e^{iv_2\eta_2}$, $S \to e^{iv\eta_s}$, then we have the axion current:

$$j_{\mu a} \sim \frac{2}{3}(v_1\partial_\mu\eta_1 + v_2\partial_\mu\eta_2 + v\partial_\mu\eta_s) = \frac{2}{3}\sqrt{v_1^2+v_2^2+v^2}\partial_\mu\eta_a, \tag{3}$$

where $\eta_a = (\sqrt{v_1^2+v_2^2+v^2}\,)^{-1}(v_1\eta_1 + v_2\eta_2 + v\eta_s)$ is the axion field and $f_a = \frac{2}{3}\sqrt{v_1^2+v_2^2+v^2} \sim \frac{2}{3}v$. The axion mass can be estimated by the standard current algebra method.

$$m_a^2 = \frac{m_{a_0}^2}{\sqrt{2}f_a^2 G_W} = \frac{9m_{a_0}^2}{4\sqrt{2}v^2 G_W}, \tag{4}$$

where $m_{a_0} \sim 50$keV. Combining this model with $N = 1$ supergravity and introducing SUSY breaking in the second term of the potential ν, the gravitino acquires a mass:

$$\frac{m_{3/2}}{M} = \frac{9}{8\sqrt{2}}K^2\frac{m_{a_0}^2}{m_a^2 G_W}. \tag{5}$$

It is a mass relation among gravitinos, axions and other fermions. The axion mass expression can be obtained from it:

$$m_a^2 = \frac{9}{8\sqrt{2}}K^2\frac{m_{A_0}^2}{G_W}\frac{M}{m_{3/2}} \approx 1.6\times 10^{-23}\frac{M}{m_{3/2}}(\text{eV})^2. \tag{6}$$

If setting $f_a = M_R \approx 10^{11} \sim 10^{12}$GeV, $m_{3/2} \approx 10^3 \sim 10^4$GeV, and $M = M_P \approx 10^{19}$GeV, then the axion mass window in our model is as follows:

$$2\times 10^{-5}\text{eV} < m_a < 4\times 10^{-4}\text{eV}, \tag{7}$$

3. Embedding in a Local SUSY GUT SO(10) Model

We can also similarly discuss the axion model in a local SUSY GUT SO(10) model. The symmtry breaking chain is

$$\mathrm{SO(10)}\otimes\mathrm{U_{PQ}(1)} \xrightarrow{M_u} \mathrm{SU_c(4)}\otimes\mathrm{SU_R(2)}\otimes\mathrm{SU_L(2)}\otimes\mathrm{U_{PQ}(1)} \xrightarrow{M_R} \mathrm{SU_c(3)}\otimes\mathrm{SU(2)}\otimes\mathrm{U(1)},$$

where $M_u \sim 10^{16}$GeV and $M_R \sim 10^{12}$GeV. $\mathrm{U_{PQ}(1)}$ is broken at M_R.

In this model, the superfields are

$S(54,0)$	$H^\beta(10,-2)$	$G^\beta(10,2)$	χ^α
$U(1,0)$	$\psi^\beta(16,1)$	$\bar{\psi}^\beta(16,-1)$	$\beta = 1,2$

where $a = 1, 2, 3$ is the family index. The first number in the paenthese is the dimension of representation and the second one is the quantum number of $U_{PQ}(1)$.

The superpotential is

$$f = \frac{1}{2}\mu\mathrm{Tr}(SS) + \frac{1}{3}h\mathrm{Tr}(SSS) + \sum_\beta cU(\bar{\psi}^\beta\psi^\beta - M_R^2) - \sum_\beta(aH^\beta SG^\beta + \frac{3}{2}aVH^\beta G^\beta)$$
$$+\sum_\beta b(\psi^1\psi^1H^\beta + \bar{\psi}^1\bar{\psi}^1G^\beta - \psi^2\psi^2H^\beta - \bar{\psi}^2\bar{\psi}^2G^\beta) + \sum_{a,b,\beta} h_{ab}^\beta H_i^\beta \chi_a^T\Gamma^i\chi_b. \tag{8}$$

The potential is

$$\nu = f_i^2 + \frac{1}{2} D^\alpha D^\alpha \tag{9}$$

where $f_i = \partial f / \partial z_i$, z_i stands for all fields. The VEV's are

$$< S >= S_0 = v(1,1,1,1,1,1,-\frac{3}{2},-\frac{3}{2},-\frac{3}{2},-\frac{3}{2}), v = \frac{2\mu}{h};$$

$$< \psi_{16}^\beta >=< \bar{\psi}_{16}^\beta >= X_0 = M_R/\sqrt{2}; \qquad < H^\beta >=< G^\beta >= U = 0. \tag{10}$$

After coupling it with $N = 1$ supergravity, the VEV's are determined by

$$f_i + \frac{1}{2} K^2 z_i^* f = 0. \tag{11}$$

Thus eq. (10) becomes

$$< S >= S_0 + S_1, S_1 = (0,0,0,0,0,0,+\varepsilon,+\varepsilon,-\varepsilon,-\varepsilon), \qquad \varepsilon = -\frac{5}{8} K^2 h v^3;$$

$$< \psi_{16}^\beta >=< \bar{\psi}_{16}^\beta >= X_0 + \varepsilon^2; < H_0^\beta >=< G_0^\beta >= P_+; < H_3^\beta >=< G_3^\beta >= iP_-; < U >= -\frac{\varepsilon}{c}, \tag{12}$$

where $P_\pm = (5Khv^2/2\sqrt{2})[(1/a)(1/2 \pm 1/h)]^{1/2}$. Choosing these coefficients appropriately, the expecting low-energy behaviour can be obtained.

4. Summary and Discussion

In both SUSY SM and SUSY GUT models, we set the $U_{PQ}(1)$ breaking scale f_a at $M_R = 10^{12}$GeV. The axion mass is around a few $\times 10^{-5}$eV. It has been shown such axions would provide closure density, and would be the dark matter. Our axion mass window eq. (7) agrees with most constraints from cosmology and astrophysics

The axion interacts with the photon in analogy to π^0. This interaction allows the axion decay to 2γ. If one photon is emitted along the direction of the axion moving and the other in opposite direction, there is a difference in the photon energy between these two directions. This provides an opportunity of searching for axions. Several optical experiments have been proposed and some of them are on the way [2]. Because $g_{a\gamma\gamma}$ is strongly sensitive to the PQ charge assignments. In our SUSY SM axion model, the PQ charges are the same as those in DFSZ models. The results of such experiments will probably distinguish different models.

The source of cosmic axions is cosmic strings. From the radiation of cosmic strings, one can calculate and the axion density and its fluctuations. It is well known that large amplitude density fluctuations produced on scales of the horizon at the QCD epoch will cause gravitational bound "miniclusters". If the axion miniclusters exist, they can be detected by femtolensing (or picolensing) because they naturally meet all three conditions of detecting [7]. This will probably provide another way to catch these elusive particles.

Acknowledgements

The author would like to thank DAAD - K.C. Wong Forschungsstipendien for the financial support and Prof. Dr. E.A. Paschos for hospitality at the Universität Dortmund, where this talk was finally completed.

References

[1] G.F. Smoot *et al., Astrophys. J.* **396** (1992) 3; S. Hancock *et al., Nature*(London) **367** (1994) 333

[2] K. van Bibber *et al.,* Preprint UCRL-JC-118357 (1994); S.L. Cheng, C.Q. Geng and W.-T. Ni, Preprint hep-ph 9506295 (1995); P.V. Vorobyov and I.V. Kolokolov, Preprint astro-ph 9501042 (1995)

[3] R. Peccei and H. Quinn, *Phys. Rev. Lett.* **38** (1977) 1440; *Phys. Rev.* **D16** (1977) 1791; S. Weinberg, *Phys. Rev. Lett.* **40** (1978) 223; F. Wilczek, *Phys. Rev. Lett.* **40** (1978) 279

[4] M. Dine, W. Fishler and M. Srednicki, *Phys. Lett.* **104B** (1981) 199; A. P. Zhitnitskii, *Sov. J. Nucl. Phys.* **31** (1980) 260

[5] J. Ellis and R. Flores, *Phys. Lett.* **263B** (1991) 259; *Nucl. Phys.* **B400** (1993) 25; *Phys. Lett.* **300B** (1993) 175.

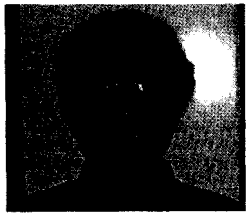
[6] K.S. Babi, K. Choi, J.C. Pati and X. Zhang, *Phys. Lett.* **333B** (1994) 364; J. Bagger, E. Poppitz and L. Randall, *Nucl. Phys.* **B426** (1994) 3; E.A. Dudas, *Phys. Rev.* **D49** (1994) 1109

[7] E.W. Kolb and I.I. Tkachev, *Phys. Rev. Lett.* **71** (1993) 3051; Preprint astro-ph 9510043 (1995)

AXIONS DE SUSY SM ET DE SUSY GUT COMME CANDIDAT À LA MATIÈRE NOIRE

RECENT RESULTS FROM THE LEP 1.5 RUN AT CERN

Eric Lançon
DAPNIA/SPP, CE-Saclay
91191 Gif/Yvette Cedex, France



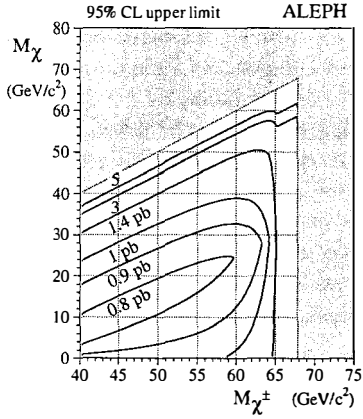
Abstract

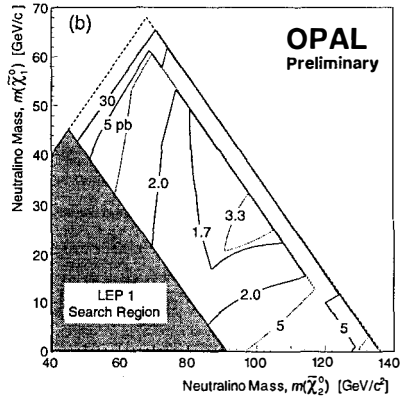
The LEP 1.5 was a three weeks period of data taking which took place in autumn 1995. The data recorded at centre-of-mass energies of 130 and 136 GeV, corresponding to a total integrated luminosity of 5.7 pb^{-1}, have been analysed by the four LEP experiments. Active searches for new particles and new phenomena have been carried out. With a limited statistic, searches for supersymmetric particles produced in e^+e^- collisions have been performed. No candidate events were found, allowing limits to be set on the masses and production cross-sections of scalar leptons, scalar tops, charginos and neutralinos. The domains previously excluded at LEP1 are substantially extended. Search for pair-produced heavy particles decaying hadronically have also been carried out in the four-jet topology. Preliminary results from the four LEP experiments[1)] are presented without any attempt to combine them.

1.0 Search for SUSY particles

Searches for SUSY particles in the framework of the Minimal Super Symmetric Model [2] (MSSM) have been performed by the four LEP experiments. Chargino pair production, is mediated by s-channel Z and photon exchange, and by t-channel sneutrino exchange. These two contributions interfere destructively, an effect which is most pronounced when the sneutrino is light and when the chargino is predominantly gaugino-like. When sneutrino exchange can be ignored, the production cross-section at 136 GeV ranges from 5.5 to 18 pb for a 60 GeV/c mass chargino, depending on its field content. The second lightest neutralino, χ^0_2, could be produced together with the lightest one by s-channel Z exchange, and by t-channel selectron exchange. In contrast to the cases of slepton and chargino production, the couplings involved in neutralino production are very strongly model dependent. Since no events have been observed, limits on production cross section can be set as shown on Figure 1 .

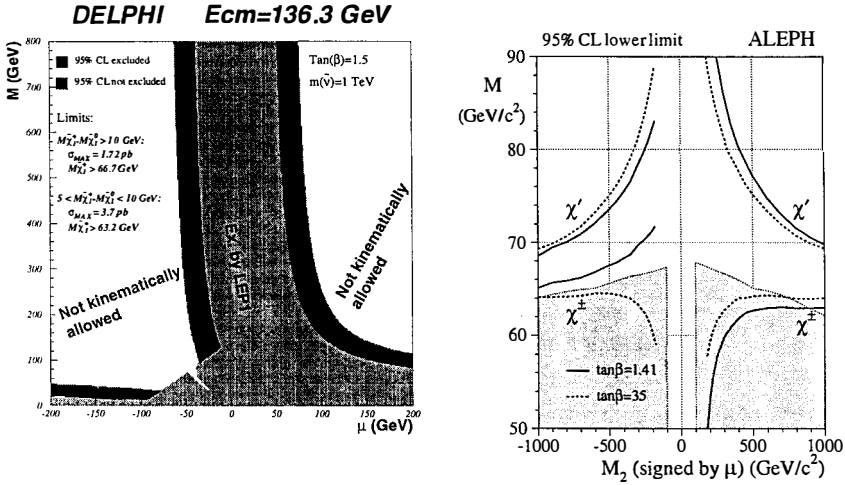
Figure 1 : **(left)** Limit on the chargino production cross section at centre-of-mass of 136 GeV, assuming W* exchange dominance in chargino decays, as a function of the chargino and neutralino masses. The lines show the upper limit at 95% confidence level. The shaded region is kinematically forbidden. **(right)** Neutralino production cross section limits.





In the MSSM, the masses and production cross-sections for charginos and neutralinos depend firstly on M_2, μ, and tanβ, and secondly on the masses of the sleptons exchanged in the t-channel (sneutrinos for charginos, selectrons for neutralinos). Consequently, the experimental results can be translated into excluded regions in the MSSM parameter space, as shown in Figure 2 (left). And lower limits on chargino and neutralino masses can be set (Figure 2 right).

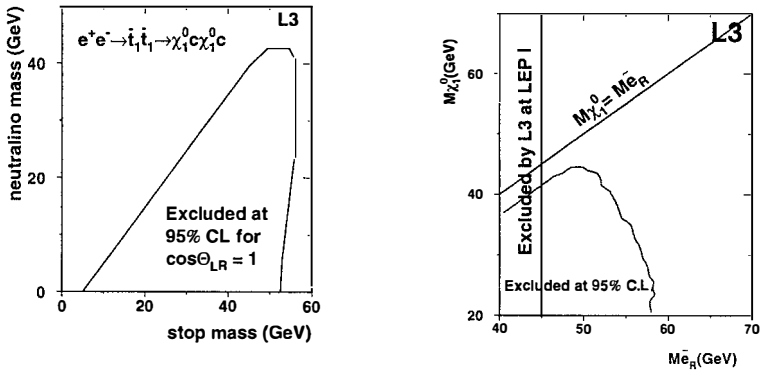
Figure 2 : **(left)** Domains in the MSSM (M_2,μ) plane excluded at LEP 1.5. **(right)** lower limit on the chargino and neutralino masses as a function of M_2. The value of M_2 is signed by μ.



The stop production cross-section depends on the mixing angle θ_{LR}. When the stop coupling to the Z vanishes *i.e.* for θ_{LR} = 0.98 rad, there is no improvement on the limits obtained at LEP1. For a purely right stop *i.e.* for cosθ_{LR}=0, the gain with respect to LEP1 is moderate: m_{stop}< 48 GeV is excluded at the 95% C.L. for m_{χ} = 30 GeV. In the theoretically unlikely case where the lightest mass eigenstate is a purely left stop *i.e.* for cosθ_{LR}=1, m_{stop} < 57 GeV is excluded at the 95% C.L. for m_{χ} <43 GeV (Figure 3 left).

For the right selectron, the production cross-section also depends on the mass and on the field content of the neutralinos which are exchanged in the t-channel. As a result, interesting limits can be obtained when these contributions are enhanced, which occurs in the ``deep gaugino'' region. In the MSSM, this corresponds to $|\mu| >> M_2$. The exclusion domain shown in Figure 3 (right) has been derived in this context, the dependence on tanβ is very weak.

Figure 3 : (left) Excluded mass region as a function of stop and neutralino masses for the choice of $\cos\theta_{LR}=1$. (right) Excluded region in the right selectron, neutralino masses plane.

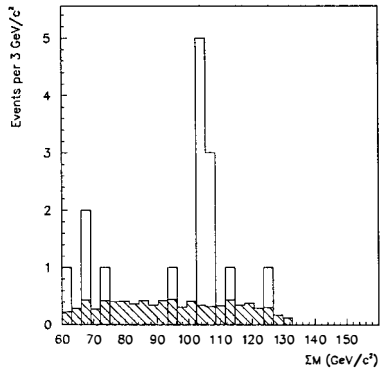


2.0 4-jet events

The analysis was first presented by the ALEPH collaboration. It was originally motivated by the search for the e^+e^- -> hA -> $b\bar{b}b\bar{b}$ process, but it can also be applied to search for other pair-produced particles. At the end of the ALEPH selection procedure, sixteen events are selected in the data while 8.6 +/- 0.3 (stat. only) were expected from the standard processes. In addition to an overall excess of events, the production of two equal (or slightly different) mass objects is also expected to be seen in the distribution of the sum ΣM of the two di-jet masses for the jet pair combination with the smallest di-jet mass difference. ΣM distribution obtained in ALEPH data is shown in Figure 4 , together with the standard model expectation. A total of twelve events is

observed above 90 GeV (where particle pair-production is not yet excluded) to be compared to the standard model expectation of 4.8 +/- 0.2 (stat. only), corresponding to a Poisson probability of 0.4%. In addition, while the background mass distribution is expected to be quite flat over the whole mass interval, the data show an accumulation of 9 events around 105 GeV. Similar analysis have been preformed by the other LEP experiments (with comparable mass resolution and efficiencies), none of the experiments find an excess (Figure 5).

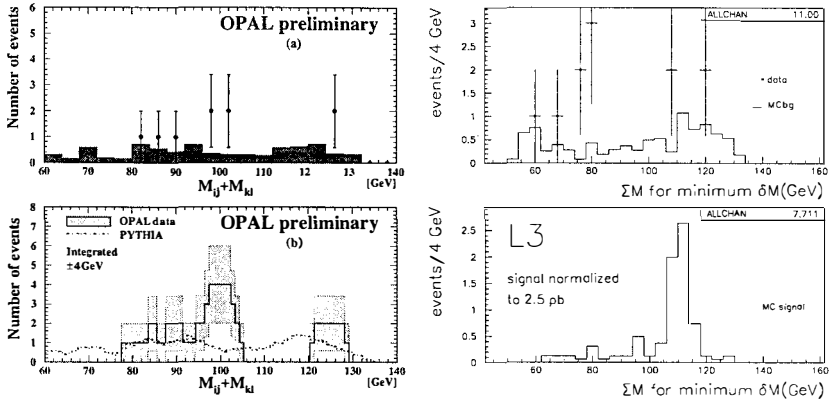
Figure 4 : Distribution of the sum of the di-jet masses for the combination with the smallest mass difference in ALEPH data. The hatched histogram is the distribution expected from a simulation of the standard processes.



3.0 Conclusions

SUSY domains previously excluded by LEP1 have substantially been extended. Kinematical limits have almost been reached. An excess of four jet events has been observed by the ALEPH experiment. Those events cluster in the di-jet mass sum distribution around 105 GeV. This excess has not been seen by the other LEP experiments. Higher statistics is therefore needed to claim the presence of new physics and, in this case, to establish a coherent scenario for it. Additional data are expected to be collected during the year 1996.

Figure 5 : **(left-a)** Sum of the two reconstructed by OPAL jet-jet masses of the events passing a selection similar to the ALEPH one. **(left-b)** The same data as (a), but plotted in a manner which would make the observation of any signal peak independent of the choice of bin boundaries. Each point shows the total number of events in a bin of two times the resolution centred at that point. **(right)** Invariant mass plot observed by L3 (top) together with the Monte-Carlo expectation. Also shown (bottom), the mass distribution for a MC signal of H^+H^- events.



Acknowledgements

The following people are greatly acknowledged for their help and cooperation in preparing this talk, Ph. Bambade, G. Cowan, L.Duflot, R. Faccini, F.F. Grivaz, P. Janot, J. Mnich, M. Schmitt, G. Taylor and D. Ward.

References

[1] **ALEPH coll.:** D. Buskulic et al., *Search for supersymmetric particles in e+e- collisions of centre-of-mass energies of 130 and 136 gev*, CERN-PPE/96-010 - submitted to Physics Letters B. **DELPHI coll.:** P. Abreu et al., *Charged particle multiplicity in e+e- interactions at sqrtS = 130 gev* - CERN-PPE/96-005 - submitted to Physics Letters B. **L3 coll.:** M. Acciarri et al., *Search for excited leptons in e+e- annihilation at sqrt{s} = 130 - 140 gev* - CERN-PPE/95-190 - submitted to Phys. Lett. B; *Measurement of hadron and lepton-pair production at 130 gev < sqrt{s} < 140 gev at lep* - CERN-PPE/95-191 - submitted to Phys. Lett. B; *Study of the structure of hadronic events and determination of alpha_s at sqrt{s}=130 gev and 136 gev* - CERN-PPE/95-192 - submitted to Phys. Lett. B. **OPAL Coll.:** Alexander et al., *Measurement of cross-sections and asymmetries in e+e- collisions at 130-140 GeV centre-of-mass energy* - CERN-PPE/96-025

[2] See for instance: H.P. Nilles, *Phys. Rep.* **110** (1984) 1; H.E. Haber and G.L. Kane, *Phys. Rep.* **117** (1985) 75; R. Barbieri, *Riv. Nuovo Cim.* **11**} no.4 (1988) 1.

STRUCTURE FORMATION AND COSMOLOGY

Observational Constraints on Ω and H_o

J.S.Bagla
Inter-University Centre for Astronomy and Astrophysics
Post Bag 4, Ganeshkhind
Pune 411 007, INDIA
E-mail : jasjeet@iucaa.ernet.in

Abstract

Here I review* some observational constraints on the cosmological parameters Ω and H_0. These constraints were chosen on the basis of reliability. Error bars have been shown explicitly to indicate the amount of freedom allowed by the constraints. I consider two classes of cosmological models, namely flat models with cosmological constant and open models without any cosmological constant. I use the standard CDM model for constraints from structure formation. Models with cosmological constant have a greater chance of surviving, as compared to the open models.

Introduction

The standard scenario in big bang cosmology assumes that at any given time the universe is homogenous and isotropic when averaged over a sufficiently large scale. The expansion of the universe is described by a scale factor "a" that satisfies the Friedmann equation

$$\left(\frac{\dot{a}}{a}\right)^2 + \frac{\kappa c^2}{a^2} = \frac{8\pi}{3} G\rho + \frac{\Lambda}{3} \quad ; \quad H_0^2 + \frac{\kappa c^2}{a_0^2} = \Omega_0 H_0^2 + \Omega_\Lambda H_0^2 \tag{1}$$

here $\kappa = 0, \pm 1$ represents curvature of the universe [$\kappa = +1$ represents a closed universe, $\kappa = -1$ an open universe and $\kappa = 0$ the flat universe.], Λ is the cosmological constant and ρ is the density of matter. The second equation is the Friedmann equation rewritten for the present epoch. $H_0 = (\dot{a}/a)_{today}$ is the Hubble's constant, Ω_0 and Ω_Λ are the densities contributed by matter and cosmological constant, respectively, in units of critical density $\rho_c = 3H_0^2/8\pi G$. Radiation does not contribute any significant amount to the energy density today, however it was the dominant constituent in the early universe and the microwave background radiation is a relic from that era.

* This work was done in collaboration with T.Padmanabhan and J.V.Narlikar[1]. Here I summarise the main results.

Structures like galaxies etc. are believed to have grown out of small inhomogenieties via gravitational instability. Observations suggest that most of the matter in the universe is invisible and this "dark" matter is likely to be nonbaryonic, noninteracting and collisionless. The only model that can be used to study generation of perturbations is the inflationary hypothesis, and this generically leads to two predictions: (i) The total density parameter $\Omega_0 + \Omega_\Lambda = 1$ and (ii) The initial power spectrum of inhomogenities has the form $P_{in}(k) \propto k^n$ with $n \simeq 1$. I will work with models that have $n = 1$. A transfer function[5] parametrised by $\Gamma \equiv \Omega_0 h$ describes changes induced in the power spectrum during linear evolution, and this is normalised with the COBE DMR observations[21] that give $Q_{rms-ps} = 18.0 \pm 1.4 \mu$K. Here Q_{rms-ps} is the inferred amplitude of fluctuations in the quadrapole.

Constraints

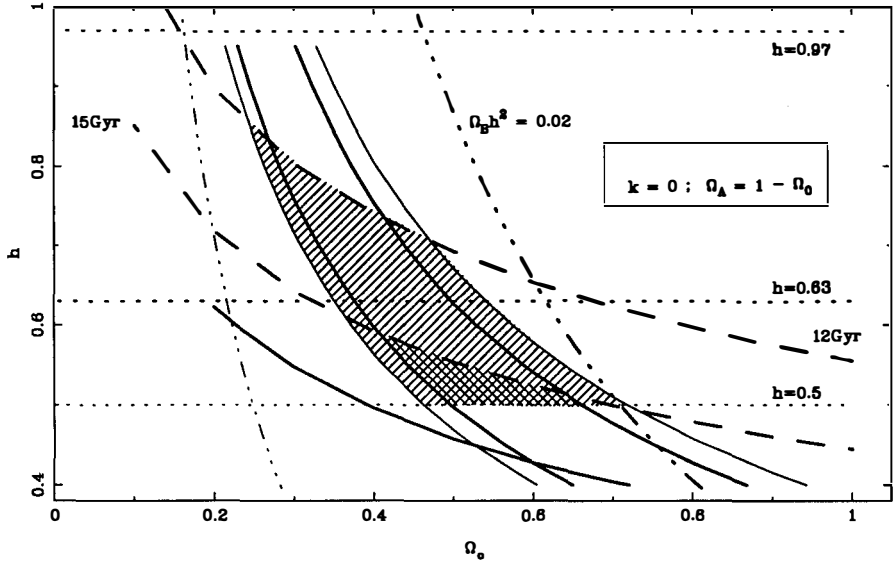
I consider constraints for two cosmological models, namely those with (i) $\Omega_0 + \Omega_\Lambda = 1$; $\kappa = 0$ and (ii) $\Omega_0 < 1$; $\Omega_\Lambda = 0$; $\kappa = -1$. These are two *natural* subsets of all models.

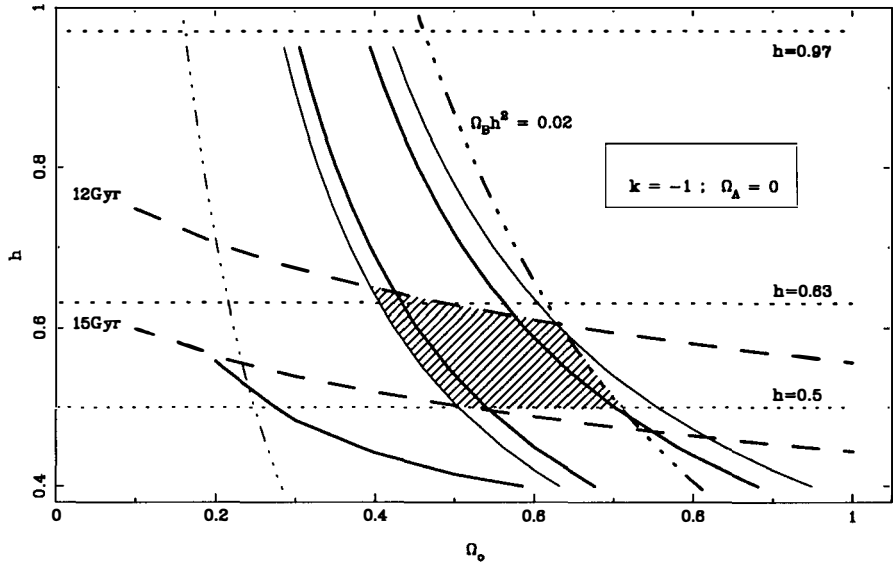
Ages of Globular Clusters : Stars in the globular clusters are the oldest known objects. By definition, the universe must be older than these. Bolte and Hogan[3] compute the ages of stars in M 92 to be 15.8 ± 2.1Gyr.

In any given model of the universe, the age can be computed and it is a function of the Hubble's constant and other cosmological parameters. Any combination of cosmological parameters that leads to an age smaller than that indicated by globular clusters is ruled out. In figure 1 I have plotted curves for $t_0 = 12$ and 15Gyr (dashed lines); the allowed region for any age lies below the corresponding curve.

Hubble's Constant : I will use the parametrization $h = H_0/100$ km s^{-1} Mpc^{-1}. To measure h, we must measure distance and recession velocity of a galaxy, or a group of galaxies. Uncertainity in measurement of recession velocity of galaxies comes mainly from their peculier motions. These can be reduced by going to large recession velocities, i.e. large distances, where the fractional error arising from peculier velocities is small. Error in the distance estimate depends upon the method that is used, and in general it increases with distance. An alternative to the standard distance indicators based on the distance ladder is a class

Figure 1 : This figure shows the constraints on the density parameter contributed by matter, Ω_0, and the Hubble's constant h arising from : (i) ages of globular clusters, (ii) measurements of Hubble's constant, (iii) abundance of rich clusters, (iv) abundance of high redshift objects, and, (v) fraction of mass contributed by baryons in clusters and primordial nucleosynthesis Top frame shows the constraints for a model with $\kappa = 0$, $\Omega_\Lambda \neq 0$ and $\Omega_0 + \Omega_\Lambda = 1$. The lower frame is for $\kappa = -1$, $\Omega_\Lambda = 0$ and $\Omega_0 \leq 1$ model. Lines of constant age are shown as dashed lines for specified values of Ω_0 and h. Dotted lines mark the band enclosing value of local Hubble constant ($0.63 < h < 0.97$) obtained from HST measurements. We have also shown the assumed lower limit for its global value ($h = 0.5$). Thick unbroken lines enclose region which is permitted by the observed abundance of clusters. Thin unbroken lines show the extent to which this region can shift due to uncertainity in the COBE normalisation of power spectrum. The almost horizontal thick unbroken line is a lower bound on permitted values of h from abundance of high redshift objects. This line depicts $\sigma(10^{11} M_\odot, z = 2) = 1$. Dot-dashed line marks the extreme upper limit allowed by primordial nucleosynthesis and fraction of mass contributed by baryons in clusters. For a given Ω_0 allowed values of h lie below this curve. Thin dot dashed line shows the upper bound implied by observation of deuterium abundance at high redshifts. Uncertainities in all observations have been included while plotting this curve. We have shaded allowed region in the parameter space by assuming that globular clusters are not older than 12Gyr and $h > 0.5$. If we take a more conservative view then the permitted area shrinks considerably and is shown as cross hatched region.


h=0.97
15Gyr
$\Omega_B h^2 = 0.02$
k = 0 ; $\Omega_\Lambda = 1 - \Omega_0$
h=0.63
12Gyr
h=0.5
h
Ω_o
0
0.2
0.4
0.6
0.8
1
0.4
0.6
0.8
1


h=0.97
$\Omega_B h^2 = 0.02$
12Gyr
k = −1 ; $\Omega_\Lambda = 0$
h=0.63
15Gyr
h=0.5
h
Ω_o
0
0.2
0.4
0.6
0.8
1
0.4
0.6
0.8
1

of "physics" based indicators which are independent of the entire distance ladder, e.g., the Sunyaev-Zeldovich effect. A reduction in errors related to modelling parameters required for analysis of this effect can lead to very accurate determination of the Hubble's constant.

Measurement[6] of distance to M 100 (a galaxy in the Virgo cluster) by the Hubble space telescope, with the use of the cepheid period luminosity relation, gives the value $h = 0.80 \pm 0.17$. This "local" value of Hubble constant may differ somewhat from its global value. Some authors[17,9] have computed the probability distribution for the difference between the global and local value. They show that values smaller than $h = 0.5$ are ruled out at 94% confidence level, given the local HST measurement. Global value of Hubble's constant can be measured directly with the Sunyaev-Zeldovich effect. Observations of Abell 2218 give[2] $h = 0.65 \pm 0.25$. Using supernovae as distance indicators leads to similar values of h [14].

In figure 1 I have plotted dotted lines bounding the region allowed by the value obtained for M 100 ($0.63 < h < 0.97$) and also for $h = 0.5$ as a lower limit for the global value of the Hubble constant.

Abundance of Rich Clusters : Mass per unit volume contained in rich clusters can be estimated from the observed number density of such clusters and their average mass. Mass of these clusters can be estimated by a variety of methods like assuming virial equilibrium and using the velocity dispersion of galaxies, gravitational lensing etc. One way of representing the observed number is to state the contribution of mass in these clusters to the density parameter, $\Omega^{obs}_{clusters}$. This number can be computed for any theoretical model using the Press-Schechter method[13] and successful models should satisfy the equality $\Omega(> M_{clusters}) = \Omega^{obs}_{clusters}$, within the errors of observations.

A comparison of observations with theory can also be carried out in a more involved manner by converting the number density of clusters into amplitude of density fluctuations at their mass scale. This amplitude is then scaled to $8h^{-1}Mpc$, a scale typically associated with cluster mass, assuming power law form for σ, the *rms* fluctuations in density perturbations. The index is chosen to match that expected in the model being considered[19]. The result is expressed as a constraint on σ_8, the rms amplitude of fluctuations at $8\mathrm{h}^{-1}$Mpc.

I have used observational constraints given by Viana & Liddle[18]. For flat model the constraints are similar to those discussed in ref.19. I have [in figure 1] plotted thick lines showing region within one sigma of the mean. Thin lines show the bounds if the uncertainity in COBE normalisation is taken into account.

Baryon content of galaxy clusters : Rich clusters of galaxies like the Coma cluster have been studied in great detail. It is possible to determine the fraction of mass contributed by baryons to rich clusters by assuming the Coma cluster to be a prototype. It is found that[20]

$$\frac{M_B}{M_{tot}} = \frac{\Omega_B}{\Omega_0} \geq 0.009 + 0.050h^{-3/2} \qquad (2)$$

with 25% uncertainity in the right hand side. This can be combined with the value of Ω_B determined from primordial nucleosynthesis to further constrain Ω_0.

Light nuclei form in the early universe as it cools from a very dense high temperature phase. Relative abundance of different elements is a function of $\Omega_B h^2$. The observed relative

abundance of elements can be used to put limits on this parameter[3,10]. I have used the values $0.01 \leq \Omega_B h^2 \leq 0.02$. By combining this value with the fraction of mass contributed by baryons in clusters we can constrain Ω_0. Plotted in figure 1 is the upper bound on matter density after the uncertainity in the observations of fraction of mass contributed by baryons has been taken into account. The permitted region lies to the left of this curve. Generalising to inhomogenous primordial nucleosynthesis does not help as that tends to reduce the value of the baryon density[8], leading to a tighter bound on Ω_0.

Recent observations[15] of Deuterium in a high redshift absorption system suggest that Ω_B is much smaller than previously thought. ($\Omega_B = 6.2 \pm 0.8 \times 10^{-3}$) This eliminates large regions from the parameter space. However, more detailed observations are required to confirm these numbers[7]. I have plotted a thin dot dashed line for the upper bound implied by this observation. The permitted region lies to the left of this curve.

Abundance of High Redshift Objects : Existence of high redshift objects like radio galaxies and damped lyman alpha systems (DLAS) allows us to conclude that the amplitude of density perturbations is of order unity at $M \simeq 10^{11} M_\odot$ at redshift $z = 2$. I have plotted this lower bound in figure 1. For flat models, the curve runs almost parallel to lines of constant age and thus provides *an upper bound for the age of the universe.* If this constraint becomes stronger or we discover globular clusters with age greater than 19Gyr, no room will be left in the parameter space we are considering. Similar results follow for open models.

A more rigorous calculation can be done along the same lines as that described for abundance of clusters. However in the case of DLAS, theoretically computed value of density parameter $\Omega(> M, z)$ should be greater than or equal to the observed value as not all systems in that mass range host a DLAS. Observations of DLAS give us mean column density $\langle \bar{N} \rangle$ of neutral hydrogen and the number of DLAS per unit redshift dN/dz. Using these and the estimated neutral fraction for gas ($f_N \sim 0.5$) we can estimate the density parameter contributed by DLAS (for more details on method of comparison with models, see ref.16).

Apart from the constraints discussed above, observations of supernovae of type Ia, which are believed to be standard candles, can be used to determine the deceleration parameter[11,12] q_0. This constraint is independent of detailed cosmogony.

Discussion

These constraints eliminate large regions from the parameter space and the surviving region shrinks further or may even disappear if observational uncertainity is reduced. I have shaded allowed regions after taking all the constraints into account. I draw the following conclusions from this figure and the preceeding discussion.

1 $\Omega_0 = 1$ [$\Lambda = 0$] model is ruled out.

2 A confirmation of low deuterium abundance at high redshifts will rule out almost all models, irrespective of the detailed cosmogony.

3 Irrespective of the value of Hubble's constant, discovery of a globular cluster with age greater that 19Gyr will create serious difficulties for CDM cosmogonies.

Acknowledgements I thank CSIR India for continued support through the SRF grant.

References

1 Bagla, J.S., Padmanabhan, T. & Narlikar, J.V., To appear in Comments on Astrophysics (1996)
2 Birkinshaw, M. & Hughes, J.P., ApJ, 420, 33 (1994)
3 Bolte, M. & Hogan, C.J., Nature, 376, 399 (1995)
4 Copi, C.J., Schramm, D.N. & Turner, M.S., Science, 267, 192 (1995)
5 Efstathiou, G., Bond, J.R. & White S.D.M., MNRAS, 258, 1p (1992)
6 Freedman, W.L. et al., Nature, 371, 757 (1994)
7 Hogan, C.J., To appear in Cosmic Abundances, ed. G. Sonneborn and S. S. Holt, PASP conference series, (1996)
8 Leonard, R.E. & Scherrer, R.J., Preprint OSU-TA-222/95 (1995)
9 Nakamura, T. & Suto, Y., Preprint UTAP - 202/95 (1995)
10 Olive, K.A. & Scully, S.T., University of Minnesota preprint UMN-TH-1341/95 (1995)
11 R.Pain, This volume (1996)
12 Perlmutter, S. et al., ApJ, 440, L41 (1995)
13 Press, W.H. & Schechter, P., ApJ, 187, 452 (1974)
14 Saha, A. et al., ApJ, 438, 8 (1995)
15 Songaila, A., Cowie, L.L., Hogan, C.J. and Rugers, M., Nature, 368, 599 (1994)
16 Subramanian, K. & Padmanabhan T., IUCAA-5/94 ; astro-ph/9402006 (1994)
17 Turner, E.L., Cen, R. & Ostriker, J.P., Astron.J. 103, 1427 (1991)
18 Viana, P.T.P. & Liddle, A.R., Sussex preprint SUSSEX-AST 95/11-1 (1995)
19 White, S.D.M., Efstathiou, G. & Frenk, C.S., MNRAS, 262, 1023 (1993)
20 White, S.D.M., Navarro, J.F., Evrard, A.E. & Frenk, C.S., Nature, 366, 429 (1993)
21 Wright, E.L. et al., COBE preprint 96-05, astro-ph/9601059 (1996)

GALAXY FORMATION : THE LINEAR THEORY

M. Lachièze-Rey
Service d'Astrophysique, CE Saclay, 91191 Gif sur Yvette

1 Introduction

The Cosmological relativistic models describe the average properties of the Universe. In particular the Friedmann – Lemaître (FL) models assume homogeneity and isotropy of the cosmic properties and of the matter distribution, which are comforted by the observations of the cosmic microwave background (CMB), at $z \approx 1000$ and the very large scale distribution of galaxies.

Galaxies, clusters, superclusters and still larger structures presently observed have not been always present. This sets the question of when and how they were formed, the general answer invoking the mechanism of *Gravitational instability*. As initial conditions, weak initial fluctuations have been created very early during the cosmic history, and have been amplified by gravitation, the more efficient interaction at large scale. Pressure is believed to have also played a role in the latest (non linear) stages of the condensation. The role of magnetic fields, which may also be important, has not been discussed very much (see, however, Coles 1992). For gravitational instability, the important quantities are the mass density $\rho(\mathbf{x})$, the density fluctuation $\delta\rho(\mathbf{x})$, or the density contrast $\delta(\mathbf{x}) = \delta\rho/\rho$.

- A first question to arise is that of the background cosmological model : any scenario for structure formation must be played on the scene designed by a peculiar FL model, among their variety. The corresponding choice requires to assign some values to the parameters H_0, Ω and Λ. Observations seem to favor presently $H_0 \approx 70 - 80$ km.sec^{-1}Mpc^{-1} and $\Omega \approx 0.1 - 0.3$ although the most popular choices for models of structure formation are rather $H_0 \approx 50$ km.sec^{-1}Mpc^{-1} and $\Omega = 1$. Nevertheless these quantities remain free parameters for the models.

- Many observations suggest that large amounts of dark matter are present in galaxies and galaxies clusters, contributing to Ω. The respective contributions Ω_B and Ω_{NB} of baryonic and non-baryonic matter remain still unknown, although they are determinant for structure formation models. The distribution of dark matter also remains a mystery. The state of the art of the models for galaxy and structure formation is far from allowing its prediction, although it should be possible in principle. When considering fluctuations, it is necessary to distinguish matter and radiation; among matter, to distinguish between baryonic and non baryonic; and, among baryonic, between visible and non visible. A phenomenological ratio $\delta_{visible}/\delta_{total}$ is introduced as an additional free parameter, the *bias parameter*, which may vary from place to place, and also with the spatial scale.

- Finally, a model must have prescribed initial conditions : when have been imprinted the initial fluctuations, and with which characteristics, among those listed below ?

The set of answers to these questions designs a *cosmogonic model*, which allows, in principle, to follow the development of the fluctuations up to the recent stages, i.e., the formation of the real observable structures. Different kinds of predictions allow to compare it with the real world.

2 A short history of the fluctuations

2.1 The main steps

- **The origin of fluctuations**

  Very weak initial fluctuations must have been imprinted very early in the Universe by a process which remains presently unknown. Their origin may be searched at the Planck time, as the final state of an hypothetical era of quantum cosmology. They may also have been created as the result of a cosmic phase transition : for instance during inflation, or from topological defects (the most popular being cosmic strings) although there are presently no convincing models to describe such processes.

- **Relativistic amplification**

  After their creation, most fluctuations experienced a first phase of amplification, which requires a relativistic approach for different reasons:

  - The Universe was dynamically dominated by electromagnetic radiation, up to the moment of matter-radiation equivalence defined by $z_{equ} \approx 10^4 \Omega$, or $t_{equ} \approx 2\ 10^{11} \Omega^{-2}$ s.
  - Baryonic matter and electromagnetic radiation were thermally and dynamically coupled up to recombination. The sound velocity was that of the radiation gas, $c_s \approx c/\sqrt{3}$.
  - The energy density ρ was very strong.
  - The horizon had a small extent, possibly larger than the size of the fluctuations themselves.

  These conditions require a relativistic treatment, in the form of perturbations to the FRW models. The relativistic calculations (see below) allow to distinguish different kinds of perturbations, which experience different kinds of evolution, as detailed below.

- **The recombination**

  The recombination of free electrons to form atoms causes the decoupling of matter and radiation. This marks the end of the primordial Universe and changes dramatically the dynamics of the fluctuations. This is also the moment where primordial anisotropies were imprinted onto the CMB. This makes very important to examine the status of the perturbations at this period, as predicted by the various models.

- **The post recombination evolution**

  At the recombination, the – still very weak – fluctuations begin a new phase of amplification. The low level of the density contrast at the first stages allows a linear analysis. But this linear phase evolves until its own disappearance, when the fluctuations have reached a level which forbids a linear approximation. Gravitational effects become then non linear, and various dissipation mechanisms will finally help the collapse and the formation of the structures that we observe now.

2.2 Jean's length and Jean's mass

The evolution of density fluctuations result from a competition between gravity, pressure, and the cosmic expansion. The Jean's theory gives some insights of this competition. It is based onto the system of linearized hydrodynamic equations, for which a Fourier transform is made. Although the initial work of Jean neglected expansion, he obtained a criterion which also applies to the expanding Universe.

Two regimes appear according to the characteristic size L of the density perturbation. One defines the *critical Jean's length*

$$\lambda_J = \frac{2\pi}{k_J} = \sqrt{\frac{\pi c_s^2}{G < \rho >}},$$

where $< \rho >$ and c_s are respectively the average density and sound velocity of the cosmic fluid. The *Jean's mass* $M_J = (\lambda_J/2)^3 \ < \rho >$ is defined as the mass contained in a sphere having a radius λ_J. Only fluctuations with a size $\lambda > \lambda_J$ can grow.

Before the recombination, the Jean's length is about equal to the size of the particle horizon, which is itself proportional to the cosmic time t. Two different definitions of the Jean's mass are used, for this period. The most useful one refers to the mass $M_m(\lambda_{hor})$ of the (non relativistic) matter contained in a sphere of radius λ_J. Since the content of the matter contained in a fluctuation remains constant, the comparison is thus easy. Note that $M_m(\lambda_{hor}) \propto t^{3/2} \propto (1+z)^{-3}$ before z_{equ}. On the other hand it is sometimes refereed to the Jean's mass as the *total* mass contained in the horizon. Since the universe is radiation dominated at this period, it is made essentially of electromagnetic radiation and increases proportionally to t.

At the recombination, the Jean's length and the Jean's mass decrease strongly. Then,

$$\lambda_J \approx 8000 \ \Omega^{-1/2} \ h^{-1} \ (1+z)^{-1} \ \mathrm{Mpc} \quad \mathrm{and} M_J \approx 5 \ 10^5 \ \Omega^{-1/2} \ h^{-1} M_\odot,$$

which remains approximately constant with time.

2.3 Different types of fluctuations

In the frame of a given cosmological model, different types of fluctuations should be distinguished. Among **metric** fluctuations one distinguishes

- *Adiabatic* (or *isentropic* fluctuations). These scalar perturbations of the density imprint fluctuations onto the metric, with an accompanying (irrotational) velocity field. They concern equally matter and radiation, with $\delta_{matter} = (3/4) \ \delta_{total}$. They are pure compressive modes, where the wave vector $\mathbf{k} \parallel \mathbf{v}$.

  If $\lambda > \lambda_J$, δ increases according to a growing mode (there is also a decreasing one). Pressure can be neglected in first approximation. The dynamics of the growth depends on the peculiar cosmological model and on the period during the cosmic evolution. It is approximately described by a power law of time, depending on the peculiar cosmological model and of the equation of state. For instance, $\delta \propto t$ during the radiation dominated period and $\delta \propto t^{2/3}$ approximately, during the matter dominated period (the relation is exact if the Universe has the critical density).

  For $\lambda < \lambda_J$ the pressure forbids the condensation: the fluctuations oscillate like acoustic waves, almost independently of gravity which can be neglected. Their amplitude remains approximately constant in time, although damping can occur, as we see below.

- *Vectors modes* describe pure velocity fluctuations: the velocity field has a pure vorticity component, in opposition with the previous case where it was purely irrotational. There is no compressive part, i.e., no accompanying density perturbation and the wave vector $\mathbf{k} \perp \mathbf{v}$. Their amplitude decreases proportionally to $1/a(t)$, the inverse of the scale factor.

- *Tensor modes* are purely relativistic. Involving no density or velocity perturbations, they correspond to gravitational waves. They are characterized by a perturbation to the metric which decreases as $1/a(t)$. The fractional perturbation to the proper distance between freely falling particles, due to this effect, decreases with the same factor (Peebles 1980).

Entropy fluctuations do not modify the total density but only the ratio $\rho_{matter}/\rho_{radiation}$, approximately equal to the entropy per baryon. They do not perturb the the metric and remain thus *isocurvature*. Since the *relative* radiation density fluctuation is small, entropy fluctuations are very near from being isothermal. Not sensitive to gravitation, they do not increase in time (only their wavelength grows withexpansion). Entropy fluctuations possibly give rise to adiabatic fluctuations at the recombination.

3 Evolution before the recombination

3.1 Relativistic amplification

In the primordial Universe, the cosmic fluid, assumed to be perfect, is described by its stress-energy tensor T_{ij}. To study gravitational instability, all quantities, like the metric tensor g_{ij}, are considered as perturbations with respect to the FL model. The pressure fluctuation δp is related to the density fluctuation $\delta\rho$ through the equation of state of the fluid. The metric fluctuation δg_{ij} is determined partially by δT_{ij}, although not completely because of the possible additional existence of gravitational waves. On the other hand, density, pressure, and velocities are not relativistic invariants and their values, which are frame dependent, have not necessarily a clear physical signification. Even the notion of "smallness" of the fluctuations is frame dependent. I will not discuss this question in details here, whose solution is either to consider jauge-independent quantities, or to work in a defined frame and interpret correctly all quantities involved. In any case it is necessary to discard the "unphysical" modes.

The scale of the fluctuation is assumed to be smaller than the radius of curvature of the Universe but, possibly, larger than the horizon. This allows to make the approximation of a flat background universe. It is usual and convenient to use the *conformal time* η defined by $d\eta = c\ dt/a(t)$ where $a(t)$ is the scale factor. In the radiation dominated universe, $a(t) = a_0\ (t/t_\bullet)^{1/2}$, which implies $\eta = \eta_0 - c\ (t\ t_0)^{1/2}/2a_0$.

The simplicity of the FL models allows to define without ambiguity linear fluctuations in a synchronous coordinate system. Under this *linear approximation*, calculations show the existence of the three modes of metric fluctuations described above.

In the early stages of the big bang models, the Universe is radiation dominated, up to the moment of matter-radiation equivalence at z_{equ}. During this period, $a(t) \propto t^{1/2}$ and the energy density of photons $\rho_r = \sigma T^4 \propto t^{-2} \propto (1+z)^4$. On the other hand the matter density $\rho_m = \Omega\ \rho_{0,crit}\ (1+z)^3$.

3.2 Adiabatic modes

The cosmic fluid has the equation of state $p = \rho c^2/3$. For adiabatic fluctuations, the peculiar velocity u is related to density fluctuations through $\delta_{matter} = 3\ \delta/4 = u/c_s$.

If $L > L_J$, there is a growing mode (I will not consider further the decaying mode) $\delta \propto t \propto (1+z)^{-2}$. It is remarkable that the associated dimensionless fluctuation of the gravitational potential $\delta\phi/c^2$ remains constant with time. In other words density fluctuation grow but metric fluctuation remain constant.

If $L < L_J$, the solution corresponds to acoustic oscillations with sound velocity $c/\sqrt{3}$. One can neglect the gravity and the amplitude δ remains constant. Dissipation effects however may play an important role.

For adiabatic fluctuations concerning baryonic matter, the photons are scattered by the free electrons, with the Thomson cross section σ_T. Their mean free path $1/\sigma_T\ n_e$ can be less than the wavelength of the perturbation. This generates a viscosity, leading to the damping of the oscillations having a scale smaller than a characteristic length L_{Silk}. To this length corresponds the "Silk mass" $M_{Silk} \approx 10^{13}\ M_\odot$. Fluctuations with $M < M_{Silk}$ suffer damping.

3.3 Other perturbations

The existence of important initial vorticity contradicts the hypotheses at the basis of the FL models. Thus vector fluctuations are not considered in the standard models. Gravitational waves may have been generated in the primordial Universe, for instance during an inflationary phase. Their amplitude decreases with expansion as $a(t)^{-1}$. They may be present at recombination, with a level comparable to density fluctuations, and thus generate anisotropies of the CMB of the same order of magnitude than those expected from density fluctuations.

Entropy perturbations, which only concern the ratio of photons to matter densities, have no dynamics and remain frozen up to the recombination. At this moment, those with mass larger than the Jean's mass at recombination, $M_{J_{rec}} \approx 5\ 10^5(\Omega h^2)^{-1/2} M_\odot$, begin to increase like adiabatic ones.

The first models for gravitational instability involved baryonic matter only. But in the last years, the possibility of fluctuations in a (possibly dominant) non baryonic component, became more and more popular, with three essential features. First, the presence of non baryonic matter allows a higher value of Ω , since the constraint due to the primordial nucleosynthesis applies to ω_B only. This allows a more efficient dynamics. Secondly, the non baryonic component is not coupled with the electromagnetic radiation. Thus, adiabatic perturbations, for instance, do not suffer the Silk damping due to the viscous interactions between electrons and photons. On the other hand, a different process, the *collisionless damping* may concern the non baryonic fluctuations, with similar consequences: all fluctuations with scale below a given limit are washed out. This suggest to distinguish three different classes among possible candidates for non baryonic matter, according to the limiting scale for collisionless damping: for *hot dark matter* (hdm), the scale is about that of superclusters, for *warm dark matter*, it is about that of galaxies; it is much smaller for for *cold dark matter* (cdm), so that collisionless damping is inefficient for this kind of models. A mixture of different components may also be present.

3.4 The statistics of the fluctuations

The density field $\rho(\mathbf{x})$ has a spatially averaged value $< \rho >$, from which is defined the density contrast

$$\delta(\mathbf{x}) = \frac{\rho(\mathbf{x}) - < \rho >}{< \rho >}.$$

A correct model of structure formation is expected to reproduce the main statistical features of the field $\delta(\mathbf{x})$. Those derive from a distribution function $P[\delta(\mathbf{x}_1), .., \delta(\mathbf{x}_N)]$, which represents the probability that the field takes the value $\delta(\mathbf{x}_1)$ at point $\mathbf{x}_1$, $\delta(\mathbf{x}_2)$ at point $\mathbf{x}_2$, etc. Many models for structure formation are constructed on the assumption that the statistic is *Gaussian*, i.e. that the distribution function is a multigaussian. Strictly speaking, this cannot be the case, since the probability that $\delta < -1$ must be strictly zero. The consequences of the departure from gaussianity and, more generally, the influence of the rare events with large values of $\mid \delta \mid$, have not been fully explored, nor the related question of the validity of the linear hypothesis.

The field $\delta(\mathbf{x})$ is usually developed in spatial Fourier components, defining the modes $\delta_\mathbf{k}$, i.e., plane waves with wave vector $\mathbf{k}$, associated to a spatial scale $L = 2\pi/k$:

$$\delta_\mathbf{k} = \frac{1}{V} \int_V d^3x\ \delta(\mathbf{x})\ \exp(i\mathbf{k}.\mathbf{x})$$

which is inverted as

$$\delta(\mathbf{x}) = \sum_{\mathbf{k}} \delta_{\mathbf{k}} \exp(i\mathbf{k}.\mathbf{x})$$

(other normalizations can be used with V and 2π factors at different places). The isotropy of the Universe implies that the averaged modes depend on the modulus $k =\mid \mathbf{k} \mid$ of $\mathbf{k}$ only, and not on its orientation. The averaged level at scale L is characterized by the power spectrum

$$P(k) = \int_{|\mathbf{k}|=k} d^3k \ \delta_k^2,$$

a kind of averaged power over all modes having a wave vector $\mathbf{k}$ of modulus $k = 2\pi/L$ (for a more complete introductory discussion, see Bertschinger,1992). The power spectrum is the Fourier transform of the autocorrelation function of the density field. When the statistics is Gaussian, the spectrum contains all statistical information. Its knowledge makes it possible, for instance, to calculate the averaged mass or density fluctuation as

$$\sigma^2 \equiv \delta(t)^2 \equiv <\mid \delta(\mathbf{x},t) \mid^2> = \sum_{\mathbf{k}} \delta_{\mathbf{k}}^2 = \frac{1}{2\pi^2} \int_0^\infty P(k) \ k^2 \ dk. \tag{1}$$

Similarly, moments of higher order are defined as

$$\sigma_n^2 \equiv \frac{V}{(2\pi)^3} \int_0^\infty P(k) \ k^{2n} \ dk. \tag{2}$$

3.5 The linear power spectrum

The spectrum evolves with time because amplification and damping of the fluctuations depend on the spatial scale. To follow this evolution, it is convenient to define a *transfer function* $\mathcal{T}(k;t)$, the ratio of the amplitude of the mode k at instant t, divided by the initial value of the mode with same (comoving) value of k. As far as the background cosmological model and the properties of the initial conditions are defined, $\mathcal{T}(k;t)$ can be calculated (see for instance Peacock, 1992, for expressions of the transfer functions corresponding to standard models). At recombination, the statistics is obtained as $P(k,t_{rec}) = \mathcal{T}(k,t_{rec}) \ P(k,t_{in})$.

Because of the simplicity of the linear dynamics, the knowledge of some characteristics of the fluctuation in their linear stage of development would give direct informations concerning the initial conditions. It is thus highly desirable to determine observationally the linear statistics, and in particular the power spectrum. Since there is no doubt that the statistics is liner at the recombination, this statistics can be estimated from that of the Cosmic Microwave Background fluctuations (see next section). Other methods involve either the spatial distribution of objects like galaxies or galaxy clusters, or the analysis of cosmic velocity fields. Conclusions concerning the linear statistics of fluctuations suffer however from three fundamental difficulties (Peacock & Dodds, 1994). (*i*) There is no real scale where we can be certain that non linear effects have not played a role and modified the linear characteristics of the statisitics. The simple act that we observe a collapsed object like a cluster of galaxies implies that non linear effects (and also dissipative processes) have played a role. (*ii*) Because of proper velocities, the observed spatial distribution of galaxies, or other objects, does not reflect their real distribution (redshift distorsion). (*iii*) The spatial distribution of galaxies, or of any type of visible objects, has different statistical properties than that of the mass. The phenomenological modelization by a bias parameter is certainly much too simple to take correctly this discrepancy into account.

Presently, there is no very convincing argument favouring a Gaussian character of the initial statistics although all observational results remain compatible with Gaussiannity (see for instance Kaiser & Peacock 1991). Many results concerning the power spectrum (which contains all information if the statistics is Gaussian), have been obtained from galaxies and clusters

samples, as well as from velocity fields catalogs. Assuming that these results express the linear dynamics imply strong biases and redshift distorsion corrections (Peacock & Dodds, 1994). This questions the validity of the linear approximation in this context, and also of the proposed corrections for non linearity.

4 The fluctuations at the recombination

The period of recombination is fundamental for the growth of structure. First, the recombination causes a dynamical decoupling between matter and radiation. Thus, large scale matter fluctuations become free to condense, without influence of the radiation pressure. This is expressed by a strong decrease of the Jean's mass and length. Secondly, the Universe is already matter dominated at the moment of recombination, so that the dynamics of the fluctuation's growth differs than from during the radiation dominated phase. Finally, matter and radiation interact for the last time at the recombination. Radiation will then propagate (almost) without interaction up to now, where it constitutes the CMB. Matter fluctuations present at the recombination have imprinted their mark on the CMB in the form of intrinsic anisotropies. Measuring those provides thus informations about fluctuations at the recombination.

4.1 Intrinsic anisotropies of the CMB

The CMB originates from the recombination. Simple geometric arguments (in curved spacetime) indicates that a proper length d at recombination is subtended by an observation angle

$$\theta = 0.95^\circ\ \Omega\ (L/100\ h^{-1}\ \mathrm{Mpc}). \tag{3}$$

This angle corresponds to a mass M such that

$$\theta = 10'\ (\Omega^2\ h)^{1/3}\ \left(\frac{M}{10^{15}\ M_\odot}\right)^{1/3}. \tag{4}$$

Galaxy-size fluctuations are thus seen under less than 1'. On the other hand the size of the particle horizon at the recombination, $L_{horizon,rec} \approx 41\ \Omega^{-1/2}\ h^{-1}$ Mpc, is seen under an angle

$$\theta_{horizon,rec} \approx 2/3\ \ (\Omega/z_{rec})^{1/2}\ \text{radians},$$

i.e., 1 or few degrees, depending on the cosmological model. This scale is comparable to that under which we see a typical supercluster. The 7° angle corresponding to the beam of the COBE satellite subtends a comoving dimension almost 1 000 h^{-1} Mpc if $\Omega = 1$. Thus (under the conventional interpretation) the size of the fluctuations seen by the DMR are much greater than the horizon scale at the recombination, and than the dimensions of all known structures.

Different kinds of temperature anisotropies are generated in the CMB by primordial fluctuations. Their relative importance depends on the scale of the fluctuations.

4.2 The Sachs-Wolfe effect

To a density fluctuation δ of dimension L is associated a fluctuation $\delta\phi$ in the gravitational potential, with the same spatial scale. Photons issuing from the corresponding zone suffer a gravitational redshift caused by this potential well. This is the Sachs-Wolfe (SW) effect. For adiabatic fluctuations, and in a flat Universe ($k = 0$), it generates a temperature anisotropy

$$\left(\frac{\Delta T}{T}\right)_{SW} = -\frac{\delta\phi}{3c^2} \approx \frac{\delta}{3}\ \left(\frac{L}{c\ t_{rec}}\right)^2 \tag{5}$$

in the CMB. This cumulative effect dominates at the larger spatial scales, beyond about $1-2°$, i.e., about 100 Mpc in comoving units.

If initial fluctuations had a (almost) Gaussian statistics, this property should have been preserved up to the recombination. The corresponding models predict a Gaussian statistics also for the CMB fluctuations. The elements $R_{ij} =< \Delta_i \Delta_j >$ of the correlation matrix of CMB anisotropies can be calculated from the power spectrum of the potential fluctuations :

$$R_{ij} = \frac{1}{9} \int d^3k \ <| \phi |^2> \ W_i(\mathbf{k}) \ W_j(\mathbf{k}), \tag{6}$$

where $W_i(\mathbf{k})$ is a window function, i.e., the Fourier transform of the selection function characterizing the observation beam. When the fluctuations have a Gaussian statistics, it is possible to calculate the autocorrelation function $C(\theta)$ of the CMB : this requires first to calculate the statistics of the gravitational potential fluctuations from the power spectrum ; and then to calculate spherical harmonics from the Fourier modes of the development of the fluctuations.

4.3 The Doppler shift

Density fluctuations generate matter motions with (weak) velocities $\mathbf{V}$, which may be calculated from the hydrodynamic continuity equation

$$\frac{\partial \delta}{\partial t} = -\nabla.\mathbf{V} \tag{7}$$

(linear approximation). The photons scattered by the moving (baryonic) matter suffer a Doppler redshift, in one direction or the other. The corresponding temperature difference can be written

$$\left(\frac{\Delta T}{T}\right)_{Doppler} \approx \frac{v}{c} \approx \delta \ \frac{L}{ct}, \tag{8}$$

where L is the spatial scale of the fluctuation, assumed to have been created without initial velocity, and t the time since this creation.

Doppler fluctuations dominate at angular scales around 1° (corresponding to about 50 Mpc, depending on the model). According to the details of the recombination, one or some Doppler peaks appear at angular sizes which depend on the cosmological parameters. In a given model for galaxy formation, the intensity of the Doppler peak(s) depends specifically on the value of Ω_B. Thus future missions of CMB observations may allow to measure these parameters with a good precision (Delabrouille 1996)

4.4 The adiabatic component

Adiabatic fluctuations, associating matter and radiation generate temperature anisotropies in the CMB,

$$\left(\frac{\Delta T}{T}\right)_{adiabatique} = \frac{1}{4} \ \delta_{radiation} = \frac{1}{3} \ \delta_{matter}, \tag{9}$$

which may be important at scales of a few Mpc, corresponding to angular scales $\leq 5'$.

However, all anisotropies with scales smaller than a few arc minutes cannot be observed because of the depth $\Delta z \approx 80$ of the last scattering surface, due to the non instantaneous character of the recombination. All fluctuations with a comoving length smaller than about $7 \ \Omega^{1/2} h^{-1}$ Mpc are washed out (a similar destruction could happen at larger scales if there was a late strong reionization).

5 The growth of fluctuations after the recombination

After having decoupled from the photons, the baryonic matter begins its post-recombination contraction, at least for fluctuations with scales larger than the new Jean's mass $M_J \approx 10^6\ M_\odot$. A Newtonian treatment is now allowed and, in a first approximation, it is convenient to neglect the pressure. After a first linear phase, the amplification will become non linear.

5.1 Equations for the self-gravitating system

The Newtonian evolution is described by the following Vlasov-Poisson equations. This provides a particular description for the matter (here I assume that all particles have the same mass m), with a distribution function $f(\mathbf{r},\mathbf{P},t)$, where $\mathbf{P} = m\mathbf{V}$ is the impulsion and $\mathbf{V}$ the velocity of a particle.

Macroscopic quantities are defined from moments : the matter density

$$\rho(\mathbf{r},t) = \int d^3P\ f(\mathbf{r},\mathbf{P},t), \tag{10}$$

and the average impulsion at point $\mathbf{r}$,

$$\bar{\mathbf{P}}(\mathbf{r},t)\ =\ \frac{1}{\rho(\mathbf{r},t)}\int d^3P\ \mathbf{P}\ f(\mathbf{r},\mathbf{P},t) = m\ \mathbf{u}, \tag{11}$$

with $\mathbf{u}(\mathbf{r},t)$ the average velocity.

The Liouville equation takes the form

$$\frac{\partial f}{\partial t} + \frac{1}{m}\ \mathbf{P}.\frac{\partial \mathbf{f}}{\partial \mathbf{r}} - \mathbf{m}\ \frac{\mathbf{d}\phi}{\mathbf{dr}}.\frac{\partial \mathbf{f}}{\partial \mathbf{P}} = \mathbf{0}. \tag{12}$$

The gravitational potential ϕ obeys the Poisson equation

$$\Delta_r \phi = 4\pi G\ \rho.$$

Usually, the gravitational potential at a point $\mathbf{r}_0$ is taken as

$$\phi(\mathbf{r}_0,t) = -G\int d\Gamma\ \frac{f(\mathbf{q},\mathbf{P},t)}{\mid \mathbf{q}-\mathbf{r}_0\mid}, \tag{13}$$

where $d\Gamma = d^3P\ d^3q$ is the phase space volume element. Note however that there may be convergence problems as $\mathbf{r}$ tends towards zero or infinity.

The kinetic and potential energies associated to one particle are respectively $T = \mathbf{P^2}/2$ and $\phi(\mathbf{r})$. However the total energy of the system is not the sum of those, but rather $\int d^3r\ d^3P\ f\ [\mathbf{P^2}/2 + \phi(\mathbf{r})/2]$, because of the self gravitating character of the system.

5.2 The fluid equations

In many circumstances, a fluid description is sufficient to follow the properties of the flow. This is the case for the cosmological solutions, and for the main aspects of the gravitational instability. The fluid description can be considered as a peculiar case of the Vlasov description, with a appropriate closure assumption and the fluid equations are obtained by taking moments of the previous system. For a perfect fluid, the velocity moments reduce to the pressure p. The evolution of cosmic matter is described by the dynamical equation

$$\frac{\partial \mathbf{u}}{\partial t} + \mathbf{u}.\nabla_r \mathbf{u} = \frac{1}{\rho}\nabla_r p - \nabla_r \phi,$$

together with the Poisson equation. The density evolves according to mass conservation

$$\frac{\partial \rho}{\partial t} = -div_r(\rho \mathbf{u}).$$

It is classical to define the Lagrangian time derivative, i.e., following the motion, as $\frac{d}{dt} = \frac{\partial}{\partial t} + \mathbf{u}.\nabla_r$. The velocity is then $\mathbf{u} = \frac{d\mathbf{r}}{dt}$, if $\mathbf{r}(t)$ is the position of a moving fluid particle. Then the equations may be written

$$\frac{d\mathbf{u}}{dt} = \frac{1}{\rho}\nabla_r p - \nabla_r \phi, \tag{14}$$

and

$$\frac{d\rho}{dt} = -\mathbf{u}.\mathrm{grad}_r(\rho). \tag{15}$$

The fluid equations may be seen as a peculiar case of the Vlasov system, with a distribution function $f(\mathbf{r}, \mathbf{P}, t) = \rho(\mathbf{r}, t)\ \delta_D[\mathbf{P} - \bar{\mathbf{P}}(\mathbf{r}, t)]$, where δ_D is the Dirac function.

A Newtonian description of the cosmological solution

All FL models can be expressed in a Newtonian form obeying these equations. For a dust universe ($p = 0$) with no cosmological constant (to consider a more general model, it will be sufficient to modify the two Friedmann equations below), the unperturbed distribution function takes the simple form

$$f_u(\mathbf{r},\ \mathbf{P}, t) = \rho_0\ a(t)^{-3}\ \delta_D[\mathbf{P} - H(t)m\mathbf{r}],$$

where $H(t) = \dot{a}(t)/a(t)$ is the Hubble parameter and the scale factor $a(t)$ obeys the Friedmann equations

$$a^2\ a'' = -\gamma/2 \tag{16}$$

and

$$a'^2 = \gamma a^{-1} - k, \tag{17}$$

where we have defined $\gamma \equiv 8\pi\ G\ \rho_0/3$. The scale factor a is normalized to its value at a time t_0, chosen as an origin and ρ_0 denotes the value of the density at t_0 ; $k = -1, 0$, or 1 is the curvature factor of space. Note that the gravitational potential, as well as the velocity field diverge at infinity, which is the price to pay if one desires to have a Newtonian description of the cosmological solution.

Since the perfect fluid assumption is at the basis of the FL cosmology, the fluid description is also perfectly adequate.

$$\rho_u(t) = \rho_0\ a(t)^{-3}, \text{with an equation of state } p_u = C\ \rho_u$$

$$\mathbf{u}(\mathbf{r}, t)_u = H\ \mathbf{r} = (a'/a)\ \mathbf{r} \tag{18}$$

$$\phi_u(\mathbf{r}, t) = 2\pi G\ \rho_u\ \mathbf{r}^2/3. \tag{19}$$

This is the unperturbed solution for gravitational instability and the perturbed quantities are defined as the difference between total quantities and unperturbed quantities.

5.3 Perturbations

The perturbed quantities are defined as $\delta\rho = \rho - \rho_u = \delta.\rho_u$, $\delta\phi = \phi - \phi_u$, $\delta\mathbf{u} = \mathbf{V} = \mathbf{u} - \mathbf{u}_u$. Introduced in the fluid equations, and taking account explicitly of the cosmological solution, they lead to a new system with a rather complicated form. This system may be conveniently simplified by using as a new variable the comoving position $\mathbf{x} = \mathbf{r}\ a(t)$. The change of variables $\{t, \mathbf{r}\} \rightarrow \{t, \mathbf{x}\}$ implies $\partial_t\ |_r = \partial_t\ |_x\ -H\ \mathbf{x}.\nabla_x$ and the velocity is $\mathbf{V} = a(t)\partial_t\mathbf{x}$. In comoving coordinates, the system becomes.

$$\frac{\partial \mathbf{V}}{\partial t} + H\ \mathbf{V} + \frac{1}{a}\mathbf{V}.\nabla\mathbf{V} = \mathbf{g}, \tag{20}$$

$$\frac{\partial \delta}{\partial t} + \frac{1}{a}\nabla.(1+\delta)\mathbf{V} = 0, \tag{21}$$

where now the gradient and divergence are with respect to $\mathbf{x}$ and the time derivatives are at constant $\mathbf{x}$. Note than $\mathbf{g} = -\frac{1}{a}\nabla_x \delta\phi$.

5.4 The linear solution

For small enough values of the density contrast $\delta << 1$, the quadratic terms can be neglected in the equations, which become very easy to solve. The solution for the growing mode can be found for instance in Peebles (1990) or in Zel'dovich & Novikov (1983), for the different cosmological models: the density contrast increases proportionally to the linear growth's rate $D(t)$, whose exact expression depends on the cosmological model. For a flat Einstein – de Sitter model, $D(t) \propto a(t) \propto t^{2/3}$. It is remarkable that the solution for the growing mode is local, which means that the time evolution at one point can be obtained from the values of the physical quantities at the same point, without spatial integration.

The validity of the linear approximation is limited. Even for a small *average* value of the density contrast, δ takes occasionally larger values at some positions, especially when the statistics is non Gaussian. Such occurrences may be rare but rare events can *a priori* influence the global development of fluctuations through non linear coupling. The further developments of the equations at second, third, ...orders (which are very tedious) suffers from the same limitations. Note also that the agreement between numerical simulations and calculations at some order does not prove the validity of these latter, since the limited dynamical range of the simulations does not allow to reproduce, precisely, the influence of the rare events.

The Zel'dovich approximation

It is classical, in hydrodynamics, to introduce a Lagrangian formalism: each particle following the flow is represented by its initial position $\mathbf{q}$, also called its Lagrangian coordinate. At time t, its position can be written $\mathbf{x}(\mathbf{q},t) = \mathbf{q} + \mathbf{w}(\mathbf{q},t)$. This defines the displacement field $\mathbf{w}$, which introduces a time dependent mapping between the Lagrangian variable $\mathbf{q}$ and the Eulerian one $\mathbf{x}(\mathbf{q},t)$, with a Jacobian matrix $J_{ij} \equiv dx_i/dq_j$, which must be non singular in order that the Lagrangian approach makes sense (see for instance Lachièze-Rey 1992b). The density contrast is calculated as

$$1 + \delta = 1/Det[J_{ij}] = 1/J \tag{22}$$

The Lagrangian time derivative $d/dt = \partial/\partial t + (d\mathbf{x}/dt).d/d\mathbf{x}$ is adapted to follow motion with Lagrangian coordinates (the dot means scalar product).

The Lagrangian formalism is exact, as far as no orbit crossing occurs. In the linear approximation, it is easy to check that the displacement vector remains proportional to its initial value, according to

$$\mathbf{w}(\mathbf{q},t) = D(t)\ \mathbf{w}_0(\mathbf{q}), \tag{23}$$

where $D(t)$ is the linear solution for the growth of fluctuations (see above). Zel'dovich has proposed to use the previous formula beyond the range of validity for the linear approximation, the density being calculated through equ.(22). This "Zel'dovich approximation" (ZA) goes beyond the linear approximation, and has been shown to give a correct description of the mildly non linear development, at least in some simple situations, in particular those corresponding to "pancakes", i.e., fluctuations for which one proper value of the Jacobian matrix dominates the two others. The ZA may for instance be used for the reconstruction of density from cosmic velocity fields, for which it is generally estimated that a mildly non linear treatment is correct (see for instance Nusser & Dekel 1992, or Lachièze-Rey 1992a).

The validity of the ZA is however also limited however, as it can be seen from the fact that the formula leads rapidly to infinite density. Various attempts have been made to extend its

validity. The most convincing one has been proposed by Gurbatov, Saichez and Shandarin (1989), but with a limited validity also. Adhesion and frozen flow approximations, Burger's equation have also been introduced in this purpose but it is very difficult to establish the (limited) range of validity of these approaches.

5.5 Hamiltonian formulation

I present here very shortly an other possible approach to the study of gravitational instability beyond the linear range. Although it has not led already to any practical result, it may help to understand some aspects of its development. In an *exterior* gravitational field, a fluid described by the Vlasov-system, may be characterized by the Hamiltonian density $\mathbf{P}^2/2m + \phi(\mathbf{r}, t)$, where $\phi(\mathbf{r}, t)$ is given from outside. To take account of the self-gravitating character of the fluid, this Hamiltonian density must be modified as

$$\mathbf{P}^2/2m + \phi(\mathbf{r}, t)/2,$$

where ϕ is explicitly obtained as a functional of the distribution function f from the solution (13) of the Poisson equation. Perez and Lachièze-Rey (1996, hereafter PLR) have shown that it is possible to take as an Hamiltonian the functional integral

$$\mathcal{H} = \int d\Gamma \; f(\mathbf{q}, \mathbf{P}, t) \; \mathbf{P}^2/2 -$$

$$G/2 \int d\Gamma \int d\Gamma' \; \frac{f(\mathbf{q}, \mathbf{P}, t) \; f(\mathbf{q}', \mathbf{P}', t)}{\mid \mathbf{q} - \mathbf{q}' \mid},$$

where $d\Gamma = d^3q \; d^3P$. They have developed an Hamiltonian theory for continuous systems with functional expressions : any physical quantity $\mathcal{A}(t)$ evolves through

$$\frac{d\mathcal{A}}{dt} = M_3[\mathcal{A}, \mathcal{H}],$$

where M_3 is a bilinear operator (a Lie bracket) which applies to $\mathcal{A}$ and $\mathcal{H}$, both considered as functionals of f.

Explicitly, this requires first to calculate the functional derivative $\frac{\delta \mathcal{A}}{\delta f}$ of $\mathcal{A}$ with respect to f. Then it is necessary to derivate this quantity with respect to the variables $\mathbf{x}$ and $\mathbf{P}$, which gives the quantities $\mathcal{A} \mid_{\mathbf{r}_i}$ and $\mathcal{A} \mid_{P_i}$. Applied to $\mathcal{H}$, these calculations give $\mathcal{H} \mid_{\mathbf{r}_i} = -g_i$ and $\mathcal{H} \mid_{P_i} = P_i$. Then, the application of the formula above leads to the fundamental evolution equation

$$\frac{d\mathcal{A}}{dt} = \int d\Gamma \; f(\mathbf{q}, \mathbf{P}, t) \; (P_i \; \mathcal{A} \mid_{\mathbf{r}_i} + g_i \; \mathcal{A} \mid_{P_i}). \tag{24}$$

Applied to f itself, this gives the Vlasov equation. More generally, this gives the time evolution of any quantity which characterizes the fluid, i.e., which can be calculated from the distribution function. The main originality of this approach, beside the self-gravitating character of the fluid, lies in its "integral" character, based on the use of functionals the distribution function. Current work is presently adaptating this approach to a fluid (rather than Vlasov) description, and to the study of the (comoving) fluctuations themselves.

References

[1] Bertschinger E., 1992, in *New insights into the Universe*, V. J. Martinez, M. Portilla & D. Sáez eds., Springer-Verlag 1992

[2] Coles P. 1992, Comments Astroph., 16, 45

[3] Delabrouille J. 1996, this conference

[4] Gurbatov S. N., Saichez A. I. & Shandarin S. F. 1989, MNRAS, 236, 385

[5] Kaiser N. & Peacock J. 1991, Ap. J., 379, 482

[6] Lachièze-Rey M. 1992a, ApJ, 407,1

[7] Lachièze-Rey M. 1992b, ApJ, 408,403

[8] Nusser T. A., and Dekel A., 1992, Ap. J., 391, 443-452

[9] Peacock J. A., 1992, in *New insights into the Universe*, V. J. Martinez, M. Portilla & D. Sáez eds., Springer-Verlag 1992

[10] Peacock J. A. & Dodds S. J. 1994, MNRAS, 267, 1020

[11] Peebles P. J. E. 1980, *The Large scale structure of the Universe*, Princeton University Press

[12] Perez J. & Lachièze-Rey M. 1996, Ap.J., in press

[13] Shandarin S. F. & Zel'dovich Ya. B. 1989, Rev. Mod. Phys. , 61, 185

[14] Zel'dovich Ya. B. 1970, A&A, 5, 84

[15] Zel'dovich Ya. B. & Novikov I. D. 1983, *The structure and evolution of the Universe*, The University of Chicago Press

Résumé

Je présente les grandes lignes du développement des fluctuations qui ont été à l'origine de la formation des galaxies et des grandes structures cosmiques. Je restreins la présentation aux premiers moments de ce développement, caractérisé par la possibilité d'appliquer une approximation linéaire qui donne son nom à la phase correspondante de la dynamique. Je décris aussi les anisotropies attendues dans le fond diffus cosmologique, qui sont engendrées durant cette phase linéaire. Enfin, j'indique très brièvement une piste possible pour examiner le développement des fluctuations ultérieur à cette phase linéaire.

THE DISTRIBUTION OF MATTER AROUND LUMINOUS GALAXIES

B. C. Bromley, R. Laflamme, M. S. Warren, & W. H. Zurek
Los Alamos National Laboratory, Los Alamos, NM, USA

Abstract

We discuss Ω_R, a measure defined as the contribution to the cosmic mass density by material within some fixed distance R of any luminous galaxy. If all the matter in the Universe were strongly correlated with galaxies, then Ω_R would rise rapidly to the standard cosmic mass density parameter Ω at small scales corresponding to the size of a typical galaxy halo. With numerical simulations we show that in both standard and low-mass CDM models only half of the mass of the Universe lies within megaparsec of galaxies with luminosity of roughly L_* or brighter. The implications of this clustering property are considerable for conventional mass measures which treat galaxies as point particles. We explore two such measures, based on the Least Action Method and the Cosmic Virial Theorem. In the former case, the method is not likely to work on scales of a typical intergalaxy spacing; however, it may perform nicely in estimating the mass of an isolated set of galaxy groups or poor clusters. In the case of the Cosmic Virial Theorem, we find that having a large fraction of the mass in the Universe located at some distance from galaxies brings in potentially severe problems of bias which can introduce large uncertainties in the estimation of Ω.

1 Introduction

The component of our research presented at the Moriond workshop by one of us (WHZ) was the viability of cold dark matter models in light of recent COBE measurements and observations of nearby cosmic large-scale structures. The bulk of this material is being published elsewhere (see [1] for an analysis of CDM and the redshift-space power spectrum which shows no incompatibility between the models and observations; also see [2] for a broader discussion). Hence we

focus on an issue which was only lightly touched upon at the workshop but which ultimately may be important to our understanding of distribution of dark matter in the Universe.

At the Heron Island Workshop on Peculiar Velocities in the summer of 1995, Jim Peebles posed the following question: How does the of cosmic mass density that is inferred by including only the mass within some distance R of a galaxy vary with R? In a system where all of the dark matter is tightly bound to luminous galaxies, this density—call it Ω_R when expressed in units of the critical density—rises rapidly to unity at distances corresponding to the typical size of a galaxy. In systems where the galaxy distribution does not account for the total mass, Ω_R may rise much more slowly, reflecting a situation where a significant amount of dark matter lies outside of galaxies, or perhaps in large, extended halos. It is clear that Peebles' question has important ramifications for observational cosmology. For example, many conventional measures of dynamical mass in large-scale structures are based on the assumption that galaxies are point particles. If Ω_R rises slowly with R then there may exist a large and uncertain distribution of dark matter which can seriously compromise these mass measures.

Here, we explore some properties of Ω_R in the context of high and low Ω cosmogonies. We consider the implications for large-scale dynamics which originally motivated Peebles to define Ω_R. Our discussion will include the Least Action Method [3, 4] and the Cosmic Virial Theorem [5].

2 The Ω_R Measure

The density measure $\Omega_R(R)$ is most easily described in algorithmic terms given above. The prescription is to determine if an infinitesimal mass element is within a distance R of any galaxy in a Fair Sample volume V; if so it is added to a running sum of mass which, after all such elements have been examined, is divided by the total volume in the sample. To relate this measure to other statistics of large-scale structure, we consider the following more formal definition:

$$\Omega_R \equiv \frac{1}{V}\left[\int_V d\vec{r}_g \rho_g(\vec{r}_g) \int_{r<R} d\vec{r}\rho(\vec{r}-\vec{r}_g) - O(\rho,\rho_g,R)\right] , \qquad (1)$$

where ρ and ρ_g are the continuous mass density and discrete galaxy distribution, respectively, and the function O represents the mass which is counted more than once by integration in spheres (denoted by subscript $r < R$) which overlap. In the limit of small R, O vanishes and Ω_R can be immediately related to the galaxy-mass pairwise correlation function ξ_{gm} by taking an ensemble average of Ω_R. In this case, Ω_R will have the form of the average cumulative mass profile of a galaxy. However, in general, the volume of integration can be of complicated topology and an ensemble average will introduce dependencies on high-order correlations between galaxies and mass ($\xi_{ggg\ldots gm}$). In a loose sense, Ω_R as a diagnostic of structure lies between the count-in-cell statistic (e.g., [6]) and the topological genus of isodensity contours [7].

We can estimate the behavior of Ω_R in various cosmological models with high-resolution N-body simulations. We present the results from three sets of numerical data corresponding

to a COBE-normalized standard cold dark matter ($100h \equiv H = 50$ km/s/Mpc; $\Omega = 1$; $\Lambda = 0$), a standard CDM model with lower normalization ($\sigma_8 = 0.74$), and a low density CDM model in a flat universe ($h = 0.8$; $\Omega = 0.2$; $\Lambda = 0.8$). In each case, 17 million particles were used to represent the dissipationless dark matter in periodic cubes of length $125h^{-1}$ Mpc to a side; the interparticle forces were determined using a treecode with $10h^{-1}$ kpc force smoothing. Galaxy halos, candidates for realistic galaxies, are identified on the basis of local density and potential fields.

Figure 1 illustrates our results. Evidently, the profile of $\Omega_R(R)$ is only modestly sensitive to changes in normalization of the primordial density field or to the cosmic mass density. There is also significant dependence on the choice of threshold mass used to identify halos. This fact is not surprising; for a given mass distribution an increase in the galaxy number density will generally steepen the Ω_R profile by increasing the volume in which the mass is counted.

A useful length scale can also be defined for the purposes of understanding dynamics of clustering: we take $R_{1/2}$ to be such that $\Omega_R(R_{1/2}) = 0.5$. This measure gives a length that roughly characterizes the failure in the assumption that the mass of the universe can be approximated by point-like galaxies. For the full halo catalogs with ~15,000 objects with "luminosity" of roughly L_* or brighter, $R_{1/2}$ is approximately 0.5–1 h^{-1} Mpc; (the COBE-normalized $\Omega = 1$ model has the smallest $R_{1/2}$ value, and the low *Omega* model has the largest value. For the top 8000 most massive objects, $R_{1/2}$ is roughly double the value for each model.

3 A Point-Mass Approximation?

The Ω_R profiles presented in §2 suggest that the approximation of the mass distribution of the Universe as point-like galaxies may be valid only when the characteristic interaction scale between galaxies is well above $R_{1/2}$; below this value a significant fraction of the mass lies outside of the luminous galaxies and may have dynamical consequences which are neglected by the point-mass approximation.

Since $R_{1/2}$ is representative of a lower limit to interaction distances for which a point-mass approximation is expected to work well, then interactions between objects larger than galaxies, such clusters or superclusters, may be less susceptible to the effects of outlying dark matter. Of course there is an obvious difficulty with this naive line of thought if the approximation is to be applied to entities which themselves consist of many galaxies. A large astrophysical object like a supercluster can generate a gravitational potential with significant high-order multipole terms; a point-mass approximation for such an object will certainly fail.

Perhaps there exists a limited range of interaction scales which is above $R_{1/2}$ yet below the limit imposed by strong multipole interactions between extended massive objects. To determine if such a range exists, we propose a way to approximate the susceptibility of a massive object to multipole interactions, based on the evolution of an object inhomogeneity with no high-order multipole moments—a spherically symmetric overdensity. The idea is that multipole

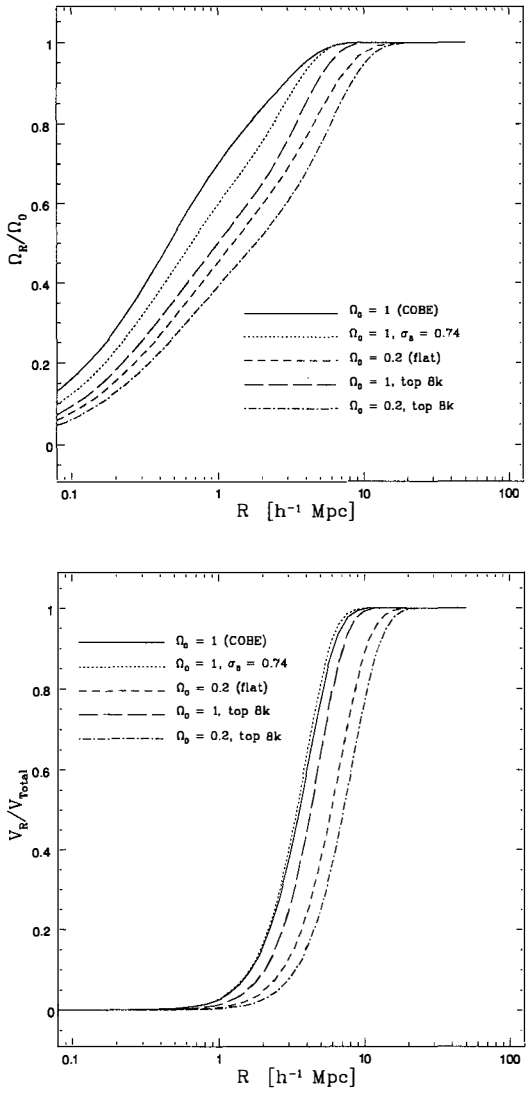


Figure 1: The profile of Ω_R, the mass within a radius R of galaxies (top) and the corresponding volume which contains that mass (bottom). Shown are profiles from standard CDM of two different normalizations and from a cosmologically flat, low density CDM model. The two rightmost curves are taken from subsets of the low-normalization standard model and the low-density model; each subset consists of the top 8,000 most massive galaxy halos. These curves shows the sensitivity of Ω_R to the galaxy mass cutoff.

interactions will be important only if an object collapses at some late time. Something which virializes at high redshift is likely to be well-approximated as a point mass, at least for the purposes of tracking interactions with similar, neighboring objects.

The spherical collapse model allows us to get a quick handle on the collapse time, here measured by z_{max}, the redshift at which an overdense region has maximum size in proper coordinates. With the definition that δ_R is the overdensity relative to the mean in a sphere of radius R, we find that

$$z_{max} \approx \begin{cases} 0.54\Omega^{0.66}\delta_R - \Omega^{-0.82} & (\Lambda = 0) \\ 0.54\Omega^{0.24}\delta_R - 0.88\Omega^{-0.17} & (\text{flat}) \end{cases} , \tag{2}$$

for cosmogonies with negligible cosmological constant Λ (upper equation) and negligible spatial curvature (lower equation). In the spirit of the Press-Schechter theory of structure formation [8] we suppose that collapsed objects originate from overdensities at the 2-σ to 3-σ level, where σ^2 is the variance in the primordial mass density field smoothed on a scale R. For galactic scales we find that for standard CDM $\delta_R \gtrsim 10$, leading to $z_{max} \gtrsim 5$, and a redshift of $z \gtrsim 2.8$ at time of collapse. This redshift estimate is reasonably large–we may at least hope that objects which reach their maximum radius at $z_{max} = 5$ have spent a good fraction of their history in a collapsed state, neither producing excessive quadrupole potential terms nor reacting strongly to tidal fields of other neighboring objects. Of course we are neglecting a great deal of physics such as merging that arises in a hierarchical clustering scenarios.

At larger scales, the fluctuation amplitudes for a 3-σ peak get considerably smaller and the redshift of maximum expansion is accordingly reduced. On a mass scale of several megaparsecs, corresponding to the Local Group, $z_{max} \approx 3$, with virialization occurring at a redshift of roughly half that value. A cluster with a mass of 10^{15} $M_\odot$ reaches its maximum radius at a redshift near unity in a standard CDM scenario. The implication is that rich clusters may have formed too late to be considered as point objects over the course of their evolution, however, groups of galaxies or poor clusters may have interacted with one another as if points since a redshift of 3 or more.

This criterion, although extremely loose, gives hope of finding dynamical systems in the universe for which the actual distribution of dark matter need not be known precisely. In this way, dynamical mass estimates may be accurately derived solely on the basis of phase-space locations of luminous objects. It remains to be determined for each such estimate whether this hope can be realized—that is, whether the set of objects for which we are looking is not empty— and this calls for detailed analysis using numerical simulations.

We now consider the implications of Ω_R for two specific measures of cosmic mass, the Least Action Method and the Cosmic Virial Theorem.

4 The Least Action Method

Peebles [3, 4] proposed that trajectories of galaxies in the Local group could be derived from an action principle operating in an expanding cosmological background spacetime with assumed cosmological parameters. An estimate of Ω then results from a best-fit between the actual phase-space locations of the galaxies (or some subset of coordinates defined in either real or redshift space) and the least action prediction. This Least Action Method (LAM), treats galaxies as point masses and its reliability hinges upon the extent to which this approximation is valid.

From Figure 1 and the discussion in §2 above, it seems unlikely that the LAM should succeed on the submegaparsec scales characteristic of intergalaxy separations within the Local Group, at least for the case of CDM cosmogonies. Simply too much mass lies outside of galaxies for the point-mass approximation to be valid on these scales. This inference from the behavior of Ω_R is actually made in retrospect. Dunn & Laflamme [9] demonstrated that that in CDM cosmogonies the LAM is seriously affected by the presence of mass outside of galaxy halos. In the context of N-body simulations, they labeled dark matter particles which are dynamically important but which are not associated with a particular galaxy halo as "orphans". Their conclusion was that the orphans cause the LAM to underestimate Ω by a significant factor.

We note that Dunn & Laflamme considered only dissipationless CDM models. The difficulties with the LAM that arise from the orphaned particles are expected to be worse in cosmogonies with hot dark matter, as the mass around galaxies is even more diffuse than for CDM. On the other hand, models with high baryonic mass fractions admit the possibility that material can cluster more densely around galaxies than in dissipationless CDM or HDM models. In this case the LAM may work well.

The LAM may also fare better when applied to objects associated with larger mass scales. As mentioned above, perhaps isolated systems of galaxy groups or poor clusters may be immune to the problems created by orphaned dark matter. We are investigating this possibility with numerical simulations although preliminary results are at best ambiguous.

5 The Cosmic Virial Theorem

Measures of kinetic and potential energy in cosmological systems are found in the pairwise radial velocity dispersion, σ_v, and the two-point correlation function, ξ_2, respectively. The Cosmic Virial Theorem (CVT) provides a means to relate these two. The mechanism is provided by the BBGKY equations which give σ_v as a function of both the two- and three-point correlation functions. Dependence on ξ_2 alone can arise only in the context of a clustering model; for example, in hierarchical models, the three-point correlation function ξ_3 has the form

$$\xi_3(\vec{x}_1, \vec{x}_2.\vec{x}_3) = Q\left[\xi_2(|\vec{x}_1 - \vec{x}_2|) + \text{cyc.}\right] \ . \tag{3}$$

Virtually all incarnations of the CVT in the literature (e.g., [5, 10, 11]) are based upon an assumption of hierarchical clustering.

If hierarchical clustering is assumed, the CVT can be expressed as

$$\sigma_v(r) = \kappa Q r^2 \xi_2(r) I(r) \Omega \tag{4}$$

where $I(r)$ has weak dependence on r, and the evaluation of the constant κ is critical to the success of the CVT as a measure of Ω.

There is a recent article by Bartlett & Blanchard [11] which discusses the CVT and the effects of using ξ_2 and σ_v as inferred from galaxies when the real issues are the velocity and spatial correlations between the galaxies and the mass in the Universe. The conclusion reached in their paper is that the distribution of matter around galaxies is extremely important for the CVT if it is based on ξ_2 and σ_v as measured from galaxies. In particular, they claimed, the low Ω value inferred from the CVT with the observed ξ_2 and σ_v for galaxies is not necessarily incompatible with an $\Omega = 1$ universe because of the presence of large, extended galaxy halos.

We can add somewhat to the discussion of Bartlett & Blanchard first by noting that the σ_v value of 350 km/s they used from the Davis & Peebles [10] analysis of the CfA I survey has been revised upward on the basis of reanalysis of the old data [12, 14, 13] and new data [15, 16]. The more recent estimates place σ_v above 700 km/s, giving a factor of 4 boost to the inferred Ω from the CVT.

However good this news may seem for those who favor a high-Ω value, caution is due to any observational estimate based on the CVT. We have worked with the CVT in numerical simulations and hold perhaps stronger views on the possible biases that can effect the CVT. Not only are there biases in σ_v and the normalization of ξ_2 between galaxies and the total matter, but Q and the power law index γ for $\xi_2 \sim r^{-\gamma}$ inferred for galaxies can be different from the mass distribution. Furthermore, even something as fundamental as the hierarchical nature of the galaxy clustering may not hold for the mass. In our simulations of structure formation in CDM we have found that all of these problems exist.

In a comparison of the simulated galaxies to the mass particles we find that (1) the galaxies have a significantly reduced σ_v profile, an effect which gets more pronounced at scales below a megaparsec; (2) the two-point correlation function of galaxies is well-fit by a power law with index $\gamma = 1.8$ on scales of $\sim$ 10 kpc to $\sim$ 10 Mpc, while ξ_2 for mass has a steeper slope, $\gamma \approx 2$, which holds down to a "core" radius of $\sim$ 100 kpc; (3) the correlation length for galaxies is $\sim$20% less than that of the mass; and (4) the Q values for mass and galaxies differ, with Q near unity for galaxies over a broad range of scales, while Q is a function of characteristic separation for mass, and rises steeply from unity below a megaparsec. This latter difference indicates that even though the galaxies may appear hierarchically clustered, the mass may not exhibit the same broad clustering property. Since the CVT as it appears in equation (4) is derived from the assumption that mass is hierarchically clustered, it is perhaps not appropriate for the CDM models. We must caution that this departure from the hierarchical clustering paradigm in the

mass distribution may be a numerical artifact, most likely the result of finite time steps in the evolution code which can wash out structure in the densest (and hottest) regions.

Despite these problems, we applied the CVT to both the mass and galaxy halo distributions in our simulations. In the case of mass, for which the CVT should work best, we found that an effective Q value near 3 gave a reasonably flat curve for Ω as a function of pair separation on scales of a megaparsec and below, suggesting that equation (4) holds despite the deviation from hierarchical clustering at small scales.

When we applied the CVT to the galaxy halos, using parameters derived from the halo distributions alone, we found that the inferred Ω was near zero for separation much above and below a megaparsec. But near the peak of the curve, at the canonical scale for applying the CVT, the inferred Ω is near unity. Unfortunately this holds true not only in the standard CDM simulations but in the model with the true Ω value of 0.2. Thus we have come up with an example, complementary to that of Bartlett & Blanchard, wherein the CVT *overestimates* Ω in a low density universe.

6 Conclusion

On the basis of a simple curve, the density of mass within a distance R of a luminous galaxy we have argued that extreme caution must be exercised in the application of mass measures that treat galaxies as point-particles. For CDM models, the distance $R_{1/2}$ which contains half of the total mass of the Universe about L_* galaxies is around a megaparsec. Below this scale the point mass approximation is likely to break down. Indeed, the numerical results for two particular mass measures, based on the Least Action Method and the Cosmic Virial Theorem, are less than promising.

With the LAM, there is hope of success when it is applied to objects larger than galaxies but much smaller than superclusters—perhaps isolated poor clusters or galaxy groups may work best. The criteria are that (1) the distance between objects to which the LAM is applied must be above $R_{1/2}$ and (2) the collapse of each individual object must have occurred at a high redshift (certainly greater than $z = 1$) so that the objects neither generate nor respond to tidal forces in a way which might significantly affect their center-of-mass motion. However the existence of such objects remains to be confirmed in simulations.

The CVT seems even more problematic because it depends on integral expressions that mix a range of scales, including those below $R_{1/2}$. The principal difficulty is that the presence of matter at some distance from galaxies may render fundamental parameters of the mass phase-space distribution (e.g., the correlation function index γ and the pairwise velocity dispersion, σ_v) irretrievable from observations of galaxies alone. However we do not reject the possibility that the CVT can be calibrated for galaxies, or perhaps a certain class of galaxies, using high-resolution numerical simulations.

References

[1] Brainerd, T. G., Bromley, B. C., Warren, M. S., & Zurek, W. H. 1996, ApJ (in press)
[2] Bromley, B. C., Brainerd, T. G., Warren, M. S., & Zurek, W. H. 1995, in Clustering in the Early Universe, Proceedings of the 1995 Moriond Meeting, ed. S. Maurogordato.
[3] Peebles, P. J. E. 1989, ApJ, 344, L53
[4] Peebles, P. J. E. 1990, ApJ, **362**, 1
[5] Peebles, P. J. E. (1980), The Large-Scale Structure of the Universe (Princeton: Princeton University Press).
[6] White, S. D. M., 1979, MNRAS, 186, 145
[7] Gott, J. R., Melott, A. L., & Dickinson, M., 1986, ApJ, 306, 341
[8] Press, W. H., & Schechter, P. 1974, *Astrophys. J.* ,, 187,425
[9] Dunn A. M., & Laflamme, R. 1995, ApJ, 443, L1
[10] Davis, M., & Peebles, P. J. E., 1983, ApJ, 267, 465
[11] Bartlett J. G., & Blanchard A., 1996, A&A, 307, 1
[12] Zurek, W. H., Warren, M. S., Quinn, P. J., & Salmon, J. K., 1993, in Proceedings of the 9th IAP Meeting, ed. Bouchet, F. R., IAP
[13] Mo, H. J., Jing, Y. P., & Börner, G., 1993, *Mon. Not. R. astr. Soc.* **264**, 825
[14] Zurek, W. H., Quinn, P. J., Salmon, J. K., & Warren, M. S. 1994, ApJ, 431, 559
[15] Marzke, R. O., Geller, M. J., da Costa, L. N., & Huchra, J. P. 1995, AJ, 110, 477
[16] Guzzo, L., Fisher, K. B., Strauss, M. S., Giovanelli, R., & Haynes, M. P., 1996, preprint (astro-ph/9504070)

LA DISTRIBUTION DE MATIERE AUTOUR DE GALAXIES LUMINEUSES

Nous allond discuter Ω_R; une mesure de la contribution a la densite de matiere cosmique de materiel a l'interieur d'une disatnce R de galaxies lumineuses. Si la matiere de l'Universe etait fortement relie aux galaxies, alors Ω_R s'accroitrait rapidement jusqu'a la densite moyenne Ω a une echelle comparable aux dimensions d'un halo typique de galaxie. Utilisant des simulations numeriques, nous demontrons que dans les models CDM standard et de faible densite, seulement la moitie de la masse totale de l'Univers se trouve dans une region d'environ 1 megaparsec du centre de des galaxies ayant une luminosite plus grande que L_*. L' importance de cette conclusion a propos de l'agregation de la matiere est considerable pour les mesures coventionels de masse qui utilise un modele ponctuel pour les galaxies. Nous explorons deux de ces mesures: la methode de moindre action (Least Action Method) et le theoreme cosmic viriel. Dans le premier cas, la methode sera probablement en difficulte a une echelle comparable a la distance entre galaxies, mais devrait etre raisonable pour estimer la masse d'un groupe isole de galaxies ou dans un agregas pauvre. Dans le deuxieme cas, nous trouvons que cette large masse qui n'appartient pas aux galaxies est problematique pour le "bias" et donc apporte une grande incertitude pour l"estimation de Ω.

Large-Scale Structure Formation in the Quasi-linear Regime

F. Bernardeau
Service de Physique Théorique, C.E. de Saclay,
F-91191 Gif-sur-Yvette, France

Abstract

The understanding of the large-scale structure formation requires the resolution of coupled nonlinear equations describing the cosmic density and velocity fields. This is a complicated problem that, for the last decade, has been essentially addressed with N-body simulations. There is however a regime, the so-called quasi-linear regime, for which the relative density fluctuations are on average below unity. It is then possible to apply Perturbation Theory techniques where the perturbation expansions are made with respect to the initial fluctuations. I review here the major results that have been obtained in this regime.

1 The Gravitational Instability Scenarios

Perturbation Theory (PT) has had important developments in the last few years simultaneously from theoretical, numerical and observational points of view. These techniques allow indeed to explore nonlinear features revealed in the statistical properties of the large-scale structures of the Universe, for which precise and robust data are now available. The general frame of these calculations is based on the dynamics of a self-gravitating *pressure-less* fluid. The large-scale structures are then assumed to have gravitationally grown from small initial fluctuations. It is important to note that in the following these initial fluctuations will be assumed to follow a Gaussian statistics. That excludes a priori exotic models that make intervene topological defects as seeds of structure formation.

By linearizing the field equations, i.e. when the quadratic coupling between the fields are neglected, one can compute the growth rate of the fluctuations. This allows to compare for instance the Cosmic Microwave Background temperature anisotropies as measured by the COBE satellite with the local density fluctuations observed in galaxy catalogues. The linear approximation cannot be used however to explore the statistical properties of the density field: the local fluctuations are just amplified, their shapes are not changed, and therefore the linear density field remains Gaussian if it obeyed such statistics initially.

2 The Perturbation Theory

The principles of these calculations have been initially presented by Peebles [35], then explored in more details in a series of recent papers [18, 22, 12, 3, 27, 28, 33, 5, 37, 38]. The starting point of all these calculations is the system of field equations, Continuity, Euler and Poisson equations, describing a single stream pressure-less fluid.

The density and velocity fields which satisfy those equations are then expanded with respect to the initial fluctuation field,

$$\delta(t,\mathbf{x}) = \sum_i \delta^{(i)}(t,\mathbf{x}), \quad \mathbf{u}(t,\mathbf{x}) = \sum_i \mathbf{u}^{(i)}(t,\mathbf{x}). \tag{1}$$

The fields $\delta^{(1)}(t,\mathbf{x})$ and $\mathbf{u}^{(1)}(t,\mathbf{x})$ are just the local linearized density and velocity fields. They are linear in the initial density field. The higher order terms, $\delta^{(2)}$, $\delta^{(3)}$,.. are respectively quadratic, cubic,.. in the initial density field, and therefore these terms do not obey a Gaussian statistics.

2.1 The Density Field

The time and space dependences of the linearized density field, $\delta^{(1)}$, factorize so that it can be written,

$$\delta^{(1)}(t,\mathbf{x}) = D(t) \int \mathrm{d}^3\mathbf{k}\, \delta(\mathbf{k}) \exp(\mathrm{i}\mathbf{k}\mathbf{x}), \tag{2}$$

where $\delta(\mathbf{k})$ are the Fourier transforms of the initial density field. They are assumed to form a set of Gaussian variables. Their statistical properties are then entirely determined by the shape of the power spectrum, $P(k)$, defined by,

$$\langle \delta(\mathbf{k})\, \delta(\mathbf{k}') \rangle = \delta_{\mathrm{Dirac}}(\mathbf{k}+\mathbf{k}')\, P(k), \tag{3}$$

where $\langle . \rangle$ denotes ensemble averages over the initial conditions. The function $D(t)$ is determined by the cosmological parameters. It is proportional to the expansion factor $a(t)$ for an Einstein-de Sitter universe. In general it has been found to depend on the cosmological density Ω in such a way that $(\mathrm{d}\log D/\mathrm{d}\log a) \approx \Omega^{0.6}$ [35] and to be very weakly dependent on the cosmological constant Λ [29].

The higher order terms can all be recursively obtained from the linear term. In general one can write the i^{th} order term as [22],

$$\delta^{(i)}(t,\mathbf{x}) = D^i(t) \int \mathrm{d}^3\mathbf{k}_1 \ldots \mathrm{d}^3\mathbf{k}_i\, \delta(\mathbf{k}_1) \ldots \delta(\mathbf{k}_i) \exp[\mathrm{i}\mathbf{x}\cdot(\mathbf{k}_1 + \ldots + \mathbf{k}_i)]\, F(\mathbf{k}_1, \ldots, \mathbf{k}_i), \tag{4}$$

where $F(\mathbf{k}_1, \ldots, \mathbf{k}_i)$ is a homogeneous function of the angles between the different wave vectors. Note that the time dependence given here is only approximate for a non Einstein-de Sitter Universe. It is however a very good approximation [3, 12, 4, 5].

2.2 The Velocity Field

For the velocity field the situation is very similar. We have first to notice that in the single flow approximation the vorticity is expected to be diluted by the expansion (eg. [35]) and thus to be negligible *at any order* of the perturbation expansion. Then it is more natural to present the properties of the velocity field in terms of the local divergence (expressed in units of the Hubble constant),

$$\theta(t,\mathbf{x}) \equiv \frac{\nabla_{\mathbf{x}} \cdot \mathbf{u}(t,\mathbf{x})}{H(t)}. \tag{5}$$

We then have,

$$\theta^{(1)}(t,\mathbf{x}) = \frac{\mathrm{d}\log D}{\mathrm{d}\log a}\ \delta^{(1)}(t,\mathbf{x}) \approx \Omega^{0.6}(t)\ \delta^{(1)}(t,\mathbf{x}). \tag{6}$$

The higher order terms can be written

$$\theta^{(i)}(t,\mathbf{x}) \approx \Omega^{0.6}(t)\ D^i(t) \int \mathrm{d}^3\mathbf{k}_1 \ldots \mathrm{d}^3\mathbf{k}_i\, \delta(\mathbf{k}_1)\ldots\delta(\mathbf{k}_i)\ \exp[\mathrm{i}\mathbf{x}\cdot(\mathbf{k}_1+\ldots+\mathbf{k}_i)]\, G(\mathbf{k}_1,\ldots,\mathbf{k}_i), \tag{7}$$

where G is another homogeneous function, different from F. Note that in general the time dependence of $\theta^{(i)}$ is not the power i of the time dependence of $\theta^{(1)}$. We will see that it induces remarkable statistical properties for the local velocity field.

2.3 Implications for the Statistical Properties of the Cosmic Fields

In general the consequences of the existence of higher order terms can be separated in two categories. These terms can affect the mean growth rate of the fluctuations or they can introduce new statistical properties because of their non-Gaussian nature. I will briefly review both aspects here.

3 The Mean Growth Rate of the Fluctuations

The mean growth rate of the fluctuation can be calculated from the shape and magnitude of the variance of the local density contrast [26, 32, 38], or equivalently of the evolved power spectrum [33, 25]. For instance the variance, $\langle\delta^2\rangle$, can be calculated using the expansion in (1),

$$\langle\delta^2\rangle = \langle\left(\delta^{(1)} + \delta^{(2)} + \ldots\right)^2\rangle. \tag{8}$$

By re-expanding this expression with respect to the initial density field one gets,

$$\langle\delta^2\rangle = \langle\left(\delta^{(1)}\right)^2\rangle + \ 2\ \langle\delta^{(1)}\ \delta^{(2)}\rangle + \langle\left(\delta^{(2)}\right)^2\rangle + \ 2\ \langle\left(\delta^{(1)}\ \delta^{(3)}\right)^2\rangle + \ldots \tag{9}$$

The term $\langle\delta^{(1)}\ \delta^{(2)}\rangle$ is zero because it is cubic in the initial Gaussian variables. The calculation of the corrective terms (the so-called "loop corrections", see [22, 3, 37, 38]), $\langle\left(\delta^{(2)}\right)^2\rangle + 2\langle\delta^{(1)}\ \delta^{(3)}\rangle$, can be done from the recursive solution of the field equations. In practice these calculations become rapidly very complicated but they can be simplified in case of a power law spectrum,

$$P(k) \propto\ k^n. \tag{10}$$

Thus for $n = -2$ one gets

$$\langle\delta^2\rangle = \langle\delta^2_{\mathrm{lin}}\rangle + \frac{1375}{1568}\ \langle\delta^2_{\mathrm{lin}}\rangle^2 \approx \langle\delta^2_{\mathrm{lin}}\rangle + 0.88\ \langle\delta^2_{\mathrm{lin}}\rangle^2 \tag{11}$$

where δ_{lin} is the linearized density contrast, $\delta_{\mathrm{lin}} = \delta^{(1)}$. This prediction can be compared with numerical results obtained from N-body codes. An interesting way to compared the two is to use the phenomenological description of the nonlinear growth rate of the fluctuations proposed by Hamilton et al. [23]. This description relates the mean linear density contrast at the scale R_{lin} to the full non-linear one at the scale $R_{\mathrm{non-lin}}$ defined by

$$R^3_{\mathrm{non-lin}}\ \left(1 + \langle\delta^2\rangle\right) = R^3_{\mathrm{lin}}. \tag{12}$$

The universal transform proposed by Hamilton et al. (given here in terms of the power spectrum) is presented in Fig. 1 (solid line) and compared to the prediction (11), (dotted line). The latter gives remarkably well the position of the departure from the pure linear regime (straight line). On the other hand the position of this transition is quite poorly given by the Zel'dovich approximation (long dashed line).

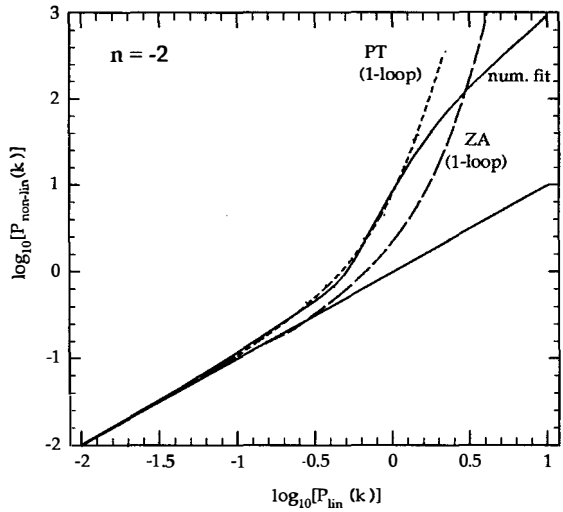


Figure 1. Comparison of the PT predictions (11) with the Hamilton et al. [23] prescription for the growth rate of the fluctuation (solid line) in case of $P(k) \propto k^{-2}$. The dotted line is the prediction of the next-to-leading order Perturbation Theory result [38], and the straight line is the linear theory prediction (figure taken from [38]).

Surprisingly the corrective terms are finite only for $n < -1$. For n larger than -1 the loop terms contain a divergence so that the results explicitly depends on a cutoff k_c introduced in the shape of the power spectrum a large k. Thus when $(R\ k_c)$ is large we have,

$$\langle \delta^2 \rangle = \langle \delta^2_{\rm lin} \rangle + C_n\ (k_c\, R)^{n+1}\ \langle \delta^2_{\rm lin} \rangle^2 \tag{13}$$

where R is the filtering radius. The existence of this divergence was not really expected since the numerical results do not indicate any significant change of behavior for $n \approx -1$. One could argue that in practice this is not a relevant problem since in realistic scenarios (like the CDM power spectrum) the index reaches $n = -3$ at small scale, so that the corrective terms are naturally regularized. However with such an interpretation it implies that the growth of the very large scale fluctuations is fed with smaller scale fluctuations, which would be slightly surprising in view of the numerical results accumulated over the last few years. Another possible explanation is that the higher order corrections induce other divergences that cancel each other (a phenomenon quite common in statistical physics). Then the corrective terms to the linear growth rate would not be proportional to the square of $\langle \delta^2_{\rm lin} \rangle$ but to a smaller power of it. Future analytic investigations may be able to throw light on this problem.

4 The Emergence of non-Gaussian Features

4.1 The Moments

The other major consequence of the existence of non-linear terms in (1) is the apparition of non-Gaussian properties. Although it is possible to characterize non-Gaussian features in many different ways, most of the efforts have been devoted to properties of the one-point probability distribution function (PDF) of the local density. More particularly, attention has been focussed on the moments if this distribution and how they are sensitive to non-linear corrections. In his treatise [35], Peebles already considered the implications of second order perturbation theory for the behavior of the third moment of the local density in case of initial Gaussian fluctuations. Indeed, using the expansion (1), the third moment reads,

$$\langle \delta^3 \rangle = \langle \left(\delta^{(1)} + \delta^{(2)} + \ldots\right)^3 \rangle = \langle \left(\delta^{(1)}\right)^3 \rangle + 3 \langle \left(\delta^{(1)}\right)^2 \delta^{(2)} \rangle + \ldots, \tag{14}$$

and what should be the dominant term of this series, $\langle \left(\delta^{(1)}\right)^3 \rangle$, is identically zero in case of Gaussian initial conditions leading to,

$$\langle \delta^3 \rangle \approx 3 \langle \left(\delta^{(1)}\right)^2 \delta^{(2)} \rangle. \tag{15}$$

Peebles calculated this expression neglecting the effects of smoothing and found (for an Einstein-de Sitter Universe),

$$\langle \delta^3 \rangle \approx \frac{34}{7} \langle \delta_{\rm lin}^2 \rangle^2, \tag{16}$$

so that the ratio, $\langle \delta^3 \rangle / \langle \delta^2 \rangle^2$, is expected to be finite at large scale. It is actually possible to extend this quantitative behavior to higher order moment and to show that the cumulants (i.e., the connected parts of the moments) are related to the second moment so that the ratios,

$$S_p = \langle \delta^p \rangle_c / \langle \delta^2 \rangle^{p-1}, \tag{17}$$

are all finite at large scale. As mentioned before the coefficient S_3 was computed by Peebles [35], Fry [18] derived S_4 and eventually Bernardeau [3] gave the whole series of these coefficients.

Unfortunately these early calculations did not take into account the filtering effects, that is that the ensemble averages should be done on the local smoothed fields. This problem was addressed numerically by Goroff et al. [22] for a Gaussian window function for the third and fourth moments. More recently these two coefficients have been calculated analytically and semi-analytically in [28, 31] for this window function. However, the results turn out to be simpler in case of a top-hat window, as it was noticed by Juszkiewicz et al. [27] for S_3 and power law spectra. The coefficients S_3 and S_4 were calculated for this window in [4] for any spectrum and any cosmological models. Eventually Bernardeau [5] proposed a method to derive the whole series of these coefficients from the spherical collapse dynamics.

I recall here the expression of the first two coefficients S_3 and S_4, as a function of the shape of the second moment,

$$S_3(R) = \frac{34}{7} + \gamma(R); \tag{18}$$

$$S_4(R) = \frac{60712}{1323} + \frac{62\gamma(R)}{3} + \frac{7\gamma(R)^2}{3} + \frac{2}{3}\frac{{\rm d}\gamma(R)}{{\rm d}\log R}; \tag{19}$$

with

$$\gamma = \frac{{\rm d}\log \sigma^2(R)}{{\rm d}\log R}. \tag{20}$$

One can notice that S_3 depends only on the local slope, and that S_4 also depends (but weakly) on the variations of that slope. These results are valid for an Einstein-de Sitter Universe but are found to be weakly Ω and Λ dependent [12, 3, 4].

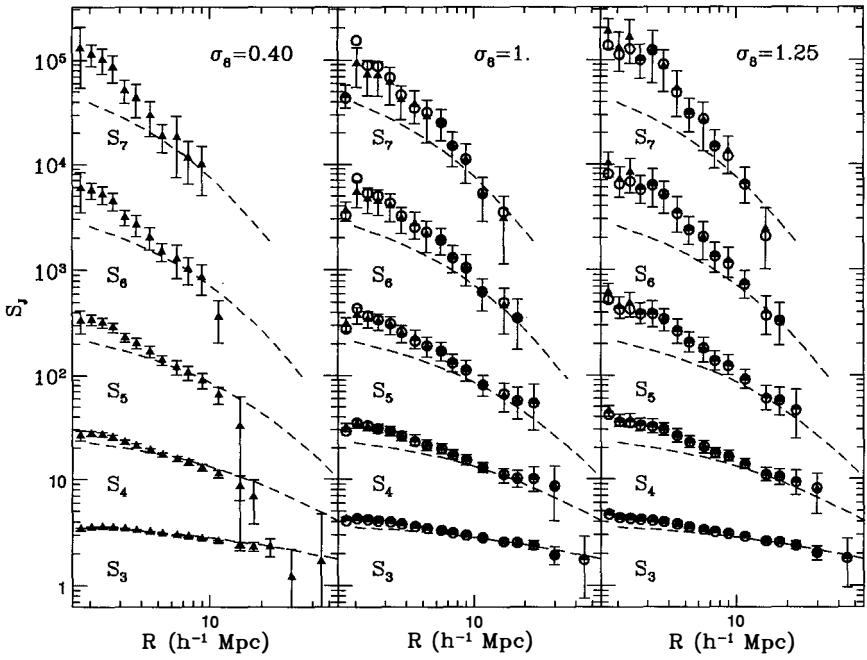


Figure 2. Comparison of PT predictions with results of numerical simulations with CDM initial conditions for moments of the local density contrast. The PT results (dashed lines) are compared to the measured S_p coefficients as a function of radius for three different time-steps (figure taken from [2]).

These coefficients have been tested against numerical results by different groups [12, 4, 2, 31] and have been found to be in extremely good agreement with the numerical results. In particular as shown in Fig.2, it has been found that the measured S_p coefficients are in good agreement with the PT results up to $p = 7$ as long as the variance $\langle \delta^2 \rangle^{1/2}$ is lower than unity [2]. This constraint does not seem to become tighter for higher orders.

4.2 From Cumulants to PDF-s

The shape of the PDF and the values of the cumulants are obviously related. When a limited number of cumulants is known it is possible to reconstruct some aspect of the shape of the PDF using the so-called Edgeworth expansion (see [28, 7]). But actually as the whole series of the cumulants is known for top-hat filtering it is possible to invert the problem (at least

numerically) and build the PDF from the generating function of the cumulants (see [1] for the general method and [3, 5] for the application to PT results). This method is justified by the fact that the S_p coefficients converge to their asymptotic PT values at roughly the same speed, i.e. for the same values of $\langle\delta^2\rangle$ (Fig. 2).

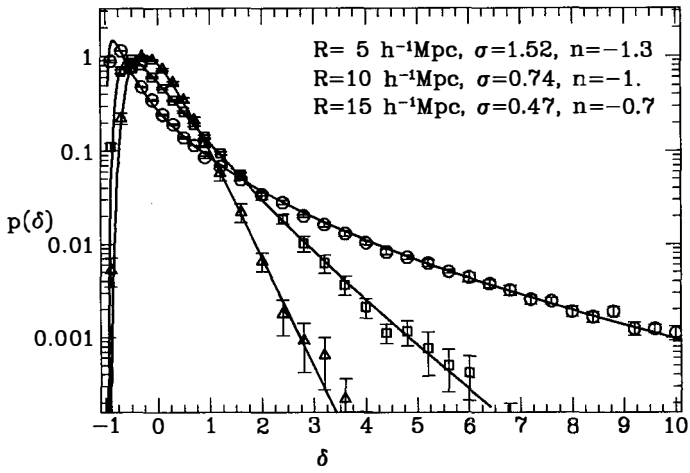


Figure 3. Comparison of PT predictions with results of numerical simulations with CDM initial conditions for the shape of the one–point PDF of the density contrast. The density PDF has been measured at three different radius, $R = 5$ h^{-1}Mpc (circles), $R = 10$ h^{-1}Mpc (squares), $R = 15$ h^{-1}Mpc (triangles) and compared to the PT predictions calculatd from the measured values of the variance and the local index (figure taken from [5]).

In Fig. 3, I present a comparison of the PT results obtained in such a way to the numerical measurements of PDF-s. Three different smoothing radii have been chosen. For each radius one can get the value of the initial index, and the value of the variance. The predicted PDF-s are entirely determined by those two numbers. As it can be seen the agreement is extremely good.

5 Comparison with data

These successes have given confidence in the validity of the PT results and have boosted investigations in the available galaxy catalogues. Using the IRAS galaxy catalogue Bouchet et al. [13] measured the 3$^{\rm rd}$ and 4$^{\rm th}$ moment of the one-point PDF of the local galaxy density. The observed relation between the third moment and the second is the following,

$$\langle\delta^3_{\rm IRAS}\rangle \approx 1.5\,\langle\delta^2_{\rm IRAS}\rangle^2. \qquad (21)$$

It is a strong indication in favor of Gaussian initial conditions since the exponent is indeed the one expected from PT. The coefficient however is lower from what is expected, but the comparison with quantitative predictions is complicated for two reasons. The galactic positions

are known only in redshift space (that is that their distance is assumed to be proportional to their line-of-sight velocity), and IRAS galaxies might be strongly biased with respect to the mass distribution. The first problem has been addressed in [24] where it is shown with adapted PT calculations that the effects of redshift space distortion on S_3 are small at large scale. It has been also investigated numerically in [30]. The fact that the measured coefficient, 1.5, is smaller than what is expected (by a factor of about 1.7) is then probably due to biases in the galaxy distribution. In case of the IRAS galaxies this is not too surprising since those galaxies are known to be under-populated (compared to bright galaxies) in very dense areas.

For these reasons a lot of interest has been devoted recently to the APM angular galaxy catalogue. The fact that it has more than 1,300,000 objects makes it the largest galaxy catalogue now available and a perfect domain of investigation. Measurements of the S_p coefficients have been made by Gaztañaga [19] for $p \leq 7$. Comparison with PT predictions are however complicated because of the projections effects. It is thus not possible to use the relation (18,19) for S_3 and S_4 and this new geometry has to be taken into account. Calculations using the small angle approximation have been made in [6] for any order of cumulants but assuming a power law spectrum. These results have been recently extended in [36] for any shape of power spectrum. In fig 4. I present the comparison of the measured S_3 and S_4 coefficients as a function of the angle (triangles) compared to the predicted ones from PT (solid lines). The latter have been calculated with either the small angle approximation or with a direct Monte-Carlo integration for which no small angle approximation is required (for S_3 only).

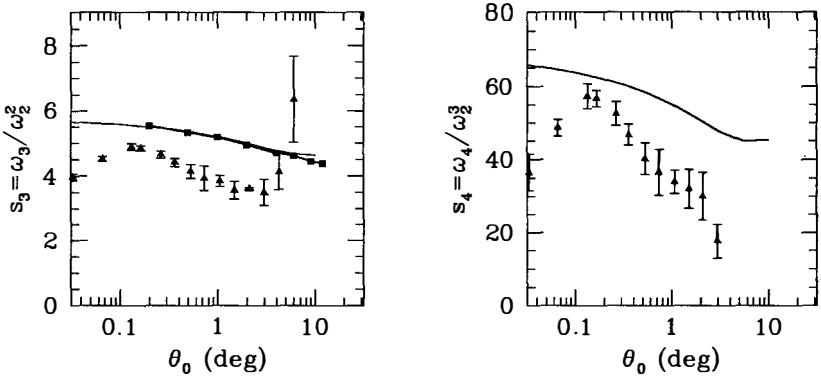


Figure 4. The $s3(\theta)$ and $s4(\theta)$ coefficients in the APM angular catalogue (triangles, see [19]) compared to the PT predictions, with the small angle approximation (solid lines), or with a numerical integration of the projection effects (squares).

There is still a significant discrepancy between the PT results and the observations. As the theoretical predictions are very robust with respect to uncertainties in the selection function or with respect to values of Ω and Λ (see [6]), it does not seem possible to reconcile the PT predictions with the observations, unless galaxies are biased. This bias is however less important than the one expected for the IRAS galaxies. These observations are anyway a strong indication in favor of Gaussian initial conditions. The implications of models with topological defects like string theories or textures have been investigated in particular by Gaztañaga [20]. He found that the current data, unless the galactic distribution is unrelated to the mass distribution(!), exclude such theories.

6 The Velocity Field Statistics

Although data on the peculiar velocities are more difficult to obtain, it is a very interesting domain of investigation since cosmic velocities are in principle directly related to the mass fluctuations. In a recent review, Dekel [15] have presented the observational and theoretical status of this rapidly evolving field.

The line-of-sight peculiar velocities are estimated from the Tully-Fisher (or similar) relation between the absolute luminosity of the galaxies and their internal velocity dispersion for elliptical, or their circular velocity for spirals. From these informations, and taking advantage of the expected absence of vorticity, it is possible to build the whole 3D smoothed velocity field. It has been done in particular in [11, 17], using the so-called POTENT reconstruction method. A straightforward application of the reconstructed velocity field is the use of the velocity-density relationship obtained in the linear regime (6) that would indeed provide a way to measure Ω. This is possible however only if galaxies are not biased. Otherwise, assuming that at large scale, $\delta^{(1)}_{\rm galaxies}(\mathbf{x}) = b\ \delta^{(1)}(\mathbf{x})$, the observations constrain a combination of Ω and b, $\beta = \Omega^{0.6}/b$. Various results for β have been given in the literature (see [15]). A rough compilation of them leads to $0.3 \lesssim \beta \lesssim 1.2$.

As we have no robust models for the large-scale bias of the galaxies, it is quite natural to explore the *intrinsic* properties of the velocity field. In the following, no assumptions are made on the bias, galaxies are simply assumed to act as *test particles* for the large-scale flows. Within this scheme Dekel & Rees [16] proposed to use the maximum expanding void to put constraints on Ω; Nusser & Dekel [34] tried to reconstruct the initial density field using the Zel'dovich approximation, thus constraining Ω on the basis of Gaussian initial conditions. Here I propose a more systematic study of the expected properties of the local divergence using PT.

In a similar way than for the density field it is indeed possible to compute the coefficients T_p that relate the high order moments of the local divergence to the second moment,

$$T_p = \langle \theta^p \rangle_c / \langle \theta^2 \rangle^{p-1}. \tag{22}$$

Unlike the S_p coefficients these ones are found to be strongly Ω dependent (but weakly Λ dependent),

$$T_p(\Omega) \propto \Omega^{-0.6(p-2)}, \tag{23}$$

as is a direct consequence of the time dependence of $\theta^{(i)}$ found in (7). The parameter $T_3(\Omega)$ was calculated in [8] as a function of Ω,

$$T_3(\Omega) = \frac{1}{\Omega^{0.6}} \left[\frac{26}{7} + (n+3) \right], \tag{24}$$

and proposed as a possible indicator to measure Ω. A preliminary investigation using the POTENT data gave $\Omega > 0.3$ with a good confidence level.

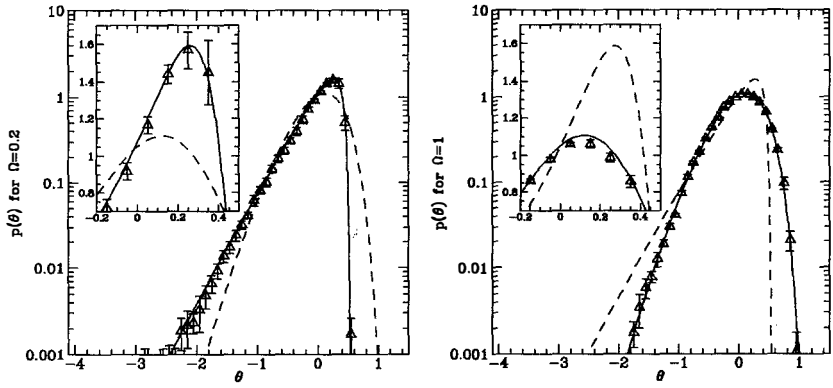


Figure 5. Comparison of the PT predictions for the shape of the PDF of the velocity divergence (solid lines) with results of numerical simulations for $n = -1$. In the left panel $\Omega = 0.2$ and in the right $\Omega = 1$. In both cases we have $\sigma_\theta \approx 0.4$. The dashed lines are the PT predictions when the reverse assumption is made on Ω.

But more generally the Ω dependence of the T_p coefficients obviously extends to the shape of the PDF of the local divergence. In particular we can show that for a power law spectrum with $n \approx -1$, it is approximately given by,

$$\lambda = 1 - \frac{2\theta}{3\Omega^{0.6}}, \quad \kappa = 1 + \frac{\theta^2}{9\lambda\Omega^{1.2}}; \tag{25}$$

$$p(\theta) = \frac{\left([2\kappa - 1]/\kappa^{1/2} + [\lambda - 1]/\lambda^{1/2}\right)^{-3/2}}{\kappa^{3/4}(2\pi)^{1/2}\sigma} \exp\left[-\frac{\theta^2}{2\lambda\sigma^2}\right], \tag{26}$$

where σ is the rms of the fluctuations of θ. This distribution exhibits a strong Ω dependence as it can be seen on Fig. 5. As expected when Ω is low the distribution is strongly non-Gaussian, with a sharp cut-off in the large divergence tails (it corresponds to rapidly expanding voids). This is this feature (discovered in an independent way) that was used in [16] to constrain Ω.

We have checked these theoretical predictions with numerical experiments (see Fig. 5) and found once again that the numerical results are in excellent agreement with them. It opens ways to have reliable measure of Ω from velocity data.

7 Conclusions

As this rapid overview have shown it, the study of the quasi-linear regime is in rapid development. A lot of efforts have been devoted to analytic calculations, and our understanding of the growth of structures in the intermediate regime has considerably improved. In particular comparisons with numerical simulations have shown that the PT results had a surprisingly large validity domain for the density as well as for the velocity fields. So far, most of the comparisons have been done for the cumulants at their leading order. However, the recent analytic results obtained for the next-to-leading order have open puzzling questions for their interpretation.

These results provide powerful tools to test the gravitational instability scenarios. Comparisons with data are extremely powerful in general. It has already provided strong indications in favor of Gaussian initial conditions. It can also give valuable indications on the existence and nature of biases for the galaxy distribution. Moreover assuming Gaussian initial conditions, it seems possible to get reliable constraints on Ω from the statistics of the large-scale cosmic flows,*independently of the biases.*

Acknowledgements. I would like to thank Roman Scoccimarro and Enrique Gaztañaga for permission to include some of their figures.

References

[1] Balian, R. & Schaeffer, R. 1989, *Astr. & Astrophys.* **220**, 1

[2] Baugh, C.M., Gaztañaga, E. & Efstathiou, G., 1994, *Mon. Not. R. astr. Soc.* in press

[3] Bernardeau, F. 1992, *Astrophys. J.* **392**, 1

[4] Bernardeau, F. 1994, *Astrophys. J.* **433**, 1

[5] Bernardeau, F. 1994, *Astr. & Astrophys.* **291**, 697

[6] Bernardeau, F. 1995, *Astr. & Astrophys.* **301**, 309

[7] Bernardeau, F. & Kofman, L., 1995, *Astrophys. J.* **443**, 479

[8] Bernardeau, F., Juszkiewicz, R., Dekel, A. & Bouchet, F. 1995, *Mon. Not. R. astr. Soc.* **274**, 20

[9] Bernardeau, F. & Van de Weygaert, R. 1996, *Mon. Not. R. astr. Soc.* **279**, 693

[10] Bernardeau, F., Van de Weygaert, R., Hivon, E. & Bouchet, F. in preparation

[11] Bertschinger, E., Dekel, A.,Faber, S.M., Dressler, A. & Burstein, D. 1990, *Astrophys. J.* **364**, 370

[12] Bouchet, F., Juszkiewicz, R., Colombi, S. & Pellat, R., 1992, *Astrophys. J.* **394**, L5

[13] Bouchet, F., Strauss, M.A., Davis, M., Fisher, K.B., Yahil, A. & Huchra, J.P. 1993, *Astrophys. J.* **417**, 36

[14] Colombi, S., Bouchet, F.R. & Schaeffer, R. 1995, *Astrophys. J. Suppl. Ser.* **96**, 401

[15] Dekel, A. 1994, *Annual Review of Astr. & Astrophys.* **32**, 371

[16] Dekel, A. & Rees, M.J. 1994 *Astrophys. J.* **422**, L1

[17] Dekel, A., Bertschinger E. & Faber, S.M. 1990, *Astrophys. J.* **364**, 349

[18] Fry, J., 1984, *Astrophys. J.* **279**, 499

[19] Gaztañaga, E., 1994, *Mon. Not. R. astr. Soc.* **268**, 913

[20] Gaztañaga, E., preprint, astro-ph/9512008

[21] Gaztañaga, E. & Frieman J. 1994, *Astrophys. J.* **437**, L13

[22] Goroff, M.H., Grinstein, B., Rey, S.-J. & Wise, M.B. 1986, *Astrophys. J.* **311**, 6

[23] Hamilton, A.J.S., Kumar, P., Lu, E. & Matthews, A. 1991 *Astrophys. J.* **374**, L1

[24] Hivon, E., Bouchet, F., Colombi, S., Juszkiewicz, R. 1995, *Astr. & Astrophys.* **298**, 643

[25] Jain, B. & Bertschinger, E. preprint, astro-ph/9503025

[26] Juszkiewicz, R. 1981, *Mon. Not. R. astr. Soc.* **197**, 931

[27] Juszkiewicz, R., Bouchet, F.R. & Colombi, S. 1993, *Astrophys. J.* **412**, L9

[28] Juszkiewicz, R., Weinberg, D. H., Amsterdamski, P., Chodorowski, M. & Bouchet, F., 1995, *Astrophys. J.* **442**, 39

[29] Lahav, O. Itoh, M. Inagaki, S. & Suto, Y. 1994 *Astrophys. J.* **402**, 387

[30] Lahav, O., Lilje, P.B., Primack, J.R. & Rees, M. 1991 *Mon. Not. R. astr. Soc.* **251**, 128

[31] Łokas, E., Juszkiewicz, R., Weinberg, D.H., Bouchet, F.R., preprint, astro-ph/9407095

[32] Łokas, E., Juszkiewicz, R., Hivon, E., Bouchet, F.R., preprint, astro-ph/9508032

[33] Makino, N, Sasaki, M., Suto, Y. 1992 *Phys. Rev. D* **46**, 585

[34] Nusser, A., Dekel, A. 1993 *Astrophys. J.* ,, 405437

[35] Peebles, P.J.E. 1980; The Large–Scale Structure of the Universe; Princeton University Press, Princeton, N.J., USA;

[36] Pollo, A. & Juszkiewicz, R. in preparation

[37] Scoccimarro, R. & Frieman, J., preprint, astro-ph/9509047

[38] Scoccimarro, R. & Frieman, J., preprint, astro-ph/9602085

Résumé

La compréhension de la formation des grandes structures requiert la résolution d'équations non-linéaires couplées décrivant l'évolution des champs de densité et de vitesse cosmologiques. C'est un problème compliqué qui, ces dix dernières années, a été traité essentiellement avec des simulations numériques à N corps. Il y a cependant un régime, le régime dit quasi-linéaire, pour lequel les fluctuations relatives de densité sont inférieures à l'unité en moyenne. Il est alors possible d'utiliser des techniques de théorie des perturbations où les développements perturbatifs sont faits par rapport aux fluctuations initiales. Je présente ici les résultats majeurs qui ont été obtenus dans ce régime.

THE STRONGLY NON-LINEAR REGIME AND N-BODY SIMULATIONS

Stéphane Colombi
CITA, University of Toronto, 60 St. George st., Toronto, Ontario, Canada M5S 3H8

Abstract

I discuss large scale structure dynamics in expanding universes filled with self-gravitating collisionless matter, paying attention to the highly nonlinear regime, which is generally addressed with N-body simulations. A statistical approach, in terms of correlation functions and moments of the probability distribution of the smoothed density field, is used. After recalling the theoretical prejudices (self-similarity, stable clustering), I discuss measurements in N-body simulations.

1 Introduction

Let us consider an expanding Friedman-Lemaître universe with zero cosmological constant filled with self-gravitating collisionless matter particles initially distributed nearly homogeneously. At the scales we are interested in (between ~ 0.1 Mpc and ~ 500 Mpc), the system is well described by the Vlasov-Poisson equations[1)]

$$\frac{\partial f}{\partial \tau} + \vec{p}.\frac{\partial f}{\partial \vec{x}} - \frac{\partial \Phi}{\partial \vec{x}}.\frac{\partial f}{\partial \vec{p}} = 0, \tag{1}$$

$$\Delta\Phi = 4\pi G a^4(\rho - \rho_b), \quad \rho = a^{-3}\int f(\vec{x},\vec{p},\tau)d^3p. \tag{2}$$

In the above equations, function f is the density in phase-space, τ is a cosmological time related to standard time by[2)] $d\tau \propto dt/a^2$, a is the expansion factor, ρ and ρ_b are respectively the spatial and the background density. The coordinates used here are the comoving position $\vec{x}$ and the momentum $\vec{p} = d\vec{x}/d\tau = a^2 d\vec{x}/dt$.

In the weakly nonlinear regime, i.e. at large scales, where the fluctuations of the density field are small, one can use an analytical, perturbative approach to solve approximately the dynamics of the system (see, e.g., Bernardeau in this book). In this paper, I discuss the nonlinear regime,

i.e. small scales, where the fluctuations of the density field are significant or large, due to structures such as sheets, filaments and clusters of galaxies. This regime is more difficult to handle analytically, mostly because of the long range nature of the gravitational forces and of the complex structure of the system in phase-space. In that case, N-body simulations are required to study the evolution of the distribution function f (except in some very peculiar cases with high level of symmetry and specific initial conditions). The problem is then to disentangle real physical processes from contamination due to numerical calculation.

There are currently two approaches to the nonlinear regime, which can correlate with each other. The first one is local. It consists in studying the evolution of particular objects, such as halos of galaxies, clusters, filaments or voids. The second approach has a global point of view. It consists in analyzing statistical properties of indicators related to the density and the velocity field. I will restrain to this approach in what follows, concentrating particularly on the N-point correlation functions of the density field and on moments of its probability distribution function, once smoothed.

This paper is thus organized as follows. In § 2, I define the statistical tools I will use. In § 3, I recall theoretical predictions, expected to apply in the highly nonlinear regime. Section 4 deals with N-body experiments. As an example, I will confront results of simulations obtained from two different codes, a high spatial resolution one (treecode), and a low spatial resolution one (PM code). Section 5 compares theoretical prejudices with measurements in N-body simulations, taking into account expected numerical contamination effects.

2 The statistical approach

The two most widely tools currently used to study large-scale structure dynamics are the correlation function $\xi_2(x)$ and its Fourier transform, the power spectrum $\langle|\delta_k|^2\rangle$ (where δ_k is the Fourier transform of the density contrast). Function $\xi_2(x)$ measures the tendancy of particles to gather by pairs. By definition, the probability δP of finding two particles respectively in two elementary volumes δV_1 and δV_2 separated by a distance x is[1)]

$$\delta P(x) \equiv n^2 \delta V_1 \delta V_2 \left[1 + \xi_2(x)\right], \tag{3}$$

where n is the number density of the considered set. Similarly, the three-point function $\xi_3(x_{12}, x_{23}, x_{31})$ measures the tendancy of particles to gather by triplets, given the fact they can already gather by pairs. By definition, the probability δP of finding three particles respectively in elementary volumes δV_1, δV_2, δV_3 at positions $\vec{x}_1$, $\vec{x}_2$ and $\vec{x}_3$ is

$$\delta P(x_{12}, x_{23}, x_{31}) \equiv n^3 \delta V_1 \delta V_2 \delta V_3 \left[1 + \xi_2(x_{12}) + \xi_2(x_{23}) + \xi_2(x_{31}) + \xi_3(x_{12}, x_{23}, x_{31})\right]. \tag{4}$$

The notation $x_{ij} \equiv |\vec{x}_i - \vec{x}_j|$ is used. One can recursively generalize the above definitions to higher order correlations functions.

In principle, the full knowledge of the statistical properties of the system requires the measurement of functions ξ_Q for any Q, except for a Gaussian distribution, for which $\xi_Q = 0$, $Q \geq 3$. The numerical estimate of function ξ_Q becomes difficult when $Q \gtrsim$ a few unities, because of the large number of parameters to handle and also because the measurements become quite noisy.

An alternate, although less complete approach, consists in concentrating on the scaling behavior of the system and forgetting about the angular dependence of the correlations. The corresponding tool is the probability distribution function $P(\rho, \ell)$ of the smoothed density field with a window of size ℓ, generally a Gaussian or a top hat filter. I assume here that the smoothing window is a spherical cell of radius ℓ and of volume $v = 4\pi\ell^3/3$. The *connected*

moments of $P(\rho,\ell)$ are nothing but the averaged correlation functions over a cell [see, e.g., refs. 3) and 4)]

$$\overline{\xi}_Q(\ell) \equiv v^{-Q} \int_v d^3x_1 \dots d^3x_Q \xi_Q(\vec{x}_1,\dots,\vec{x}_Q). \tag{5}$$

The function $P(\rho,\ell)$ is much easier to measure than the Q-point correlation functions, but involves of course much less information.

3 Theoretical prejudices[1),5)]

3.1 Similarity solution

One can search for silimilarity solutions of the Vlasov-Poisson system (1), (2) of the form $f(\vec{x},\vec{p},\tau) = \tau^u \hat{f}(\vec{x}/\tau^v, \vec{p}/\tau^w)$. This works only in the case $a \propto t^{2/3}$ ($\Omega = 1$). The solution writes

$$f(\vec{x},\vec{p},t) = t^{-3\alpha-1} \hat{f}\left(\frac{\vec{x}}{t^\alpha}, \frac{\vec{p}}{t^{\alpha+1/3}}\right). \tag{6}$$

The two-point function can be written as the following ensemble average

$$\xi_2(x,t) = \langle \left[a^{-3}\rho_b^{-1} \int d^3p_1 f(\vec{y},\vec{p}_1,t) - 1\right] \left[a^{-3}\rho_b^{-1} \int d^3p_2 f(\vec{y}+\vec{x},\vec{p}_2,t) - 1\right] \rangle. \tag{7}$$

If the similar solution (6) applies, it thus reduces to

$$\xi_2(x,t) = \hat{\xi}_2\left(\frac{x}{t^\alpha}\right). \tag{8}$$

Similarly, one can write

$$\xi_3(x_{12},x_{23},x_{31},t) = \hat{\xi}_3\left(\frac{x_{12}}{t^\alpha}, \frac{x_{23}}{t^\alpha}, \frac{x_{31}}{t^\alpha}\right). \tag{9}$$

If the similarity solution applies, one can use linear perturbation theory to relate α to initial conditions, which have then to be necessarily scale invariant. Indeed, in the linear regime, the correlation function writes $\xi_2(x,t) = a^2 \check{\xi}_2(x) = \hat{\xi}_2(x/t^\alpha)$, implying $\xi_2(x,t) \propto x^{-(n+3)}$ with

$$\alpha = \frac{4}{9+3n}. \tag{10}$$

The (initial) linear power spectrum is thus necessarily of the form

$$\langle |\delta_k|^2 \rangle_{\rm linear} \propto k^n. \tag{11}$$

3.2 Stable clustering hypothesis

Once formed, dense objects, such as halos of galaxies, tend to virialize and to keep approximately constant physical size. This can be translated in terms of "local statistical equilibrium". Then, the conditional probability of finding a particle in an elementary ***physical*** volume δV_2 at ***physical*** distance $r = ax$ from a fixed particle does not depend on time:

$$\delta P(2|1) \simeq n\,[1+\xi_2(x,t)]\,\delta V_2 = \delta P(ax). \tag{12}$$

The number density scales as $n \propto a^{-3}$, thus

$$1 + \xi_2(x,t) = a^3 F(ax). \tag{13}$$

Similarly one can show, with the same arguments

$$[1 + \xi_2(x_{12}) + \xi_2(x_{13}) + \xi_2(x_{31}) + \xi_3(x_{12},x_{23},x_{31},t)] = a^6 G(ax_{12}, ax_{23}, ax_{31}), \tag{14}$$

and so on.

3.3 Similarity plus stable clustering

Combining equations (8) and (13), we find, in the regime $\xi_2 \gg 1$, that

$$\xi_2(x,t) \propto x^{-\gamma}, \quad \gamma = \frac{6}{2+3\alpha} = \frac{9+3n}{5+n}. \tag{15}$$

Similarly, one finds that $\xi_3(\lambda x_{12}, \lambda x_{23}, \lambda x_{31}) = \lambda^{-2\gamma}\xi_3(x_{12}, x_{23}, x_{31})$. Actually, this property can be generalized at arbitrary order as the *scaling relation*[3]

$$\xi_Q(\lambda\vec{x}_1, \ldots, \lambda\vec{x}_Q) = \lambda^{-(Q-1)\gamma}\xi_Q(\vec{x}_1, \ldots, \vec{x}_Q), \tag{16}$$

which implies that the ratios

$$S_Q \equiv \frac{\bar{\xi}_Q}{\bar{\xi}_2^{Q-1}}, \quad Q \geq 3, \tag{17}$$

do not depend on scale. These results are valid only for flat universe, i.e. $\Omega = 1$, since the similarity solution does not apply otherwise. There are arguments in favor of their generalisation to the case where the density parameter Ω is not too different from unity[6].

4 N-body experiments : PM versus TREE

Most of the numerical codes which try to solve the set of equations (1) and (2) are particle based. They may be separated into two families: low *spatial resolution* (LSR) and high spatial resolution (HSR) codes. The first family is mostly represented by the PM (Particle-Mesh) codes[7]. These codes interpolate the mass of each particle on a grid to compute the density and then generally use fast Fourier transform to solve the Poisson equation. The spatial resolution of PM codes is thus given by the grid cell size, $\Delta x = L_{\rm box}/N_{\rm grid}$, with $N_{\rm grid} \sim 100 - 500$, and where $L_{\rm box}$ is the simulation box size. The low CPU cost of such codes allows a good *mass resolution*, i.e. a large number of particles, typically $N_{\rm par} \sim 10^6$ to 10^8. The second family of codes includes for example P^3M (Particle-Particle Particle-Mesh) codes[7),8] and treecodes. P^3M codes are modified PM codes. Spatial resolution is improved by computing the individual interactions between nearby particles. Treecodes decompose hierarchically the system on a tree structure, which can be for example a mutually nearest neighbor binary tree[9], or a space balanced oct tree in which each branch is a cubical portion of space[10]. In HSR codes, a softening parameter ε is introduced to bound the forces and reduce two-body (or more) relaxation effects. This parameter is generally chosen to be a small fraction of the mean interparticle distance d, typically $\varepsilon = d/(10 \text{ to } 20)$. In current HSR simulations, the number of particles, $N_{\rm par}$, is of order of 10^5 to 10^7. Spatial resolution, given by ε, is typically one order of magnitude better than in LSR codes.

As an illustrative example, I compare results from a PM simulation[11] and a treecode simulation[12] with same initial scale-free conditions $\langle|\delta_k|^2\rangle_{\rm linear} \propto k^{-1}$ (normalized to white noise at the Nyquist frequency of the particles). Both simulations assume $\Omega = 1$, involve $N_{\rm par} = 64^3$ particles in a cubical box of size $L_{\rm box}$ with periodic boundaries. The spatial resolution was $\Delta x = L_{\rm box}/128$ and $\varepsilon = L_{\rm box}/1280$ respectively for the PM and the treecode simulations. The two upper panels of Figure 1 show thin slices extracted from the simulations, after 6.4 expansion factors. The left slice corresponds to PM and the right one to tree. There are only tiny differences in general, but significant in overdense regions. One can indeed notice that clusters of particles look slightly thicker in the PM case than in the tree case. More quantitative comparisons can be made on pairwise statistics. Left bottom panel shows the quantity ξ_2 as

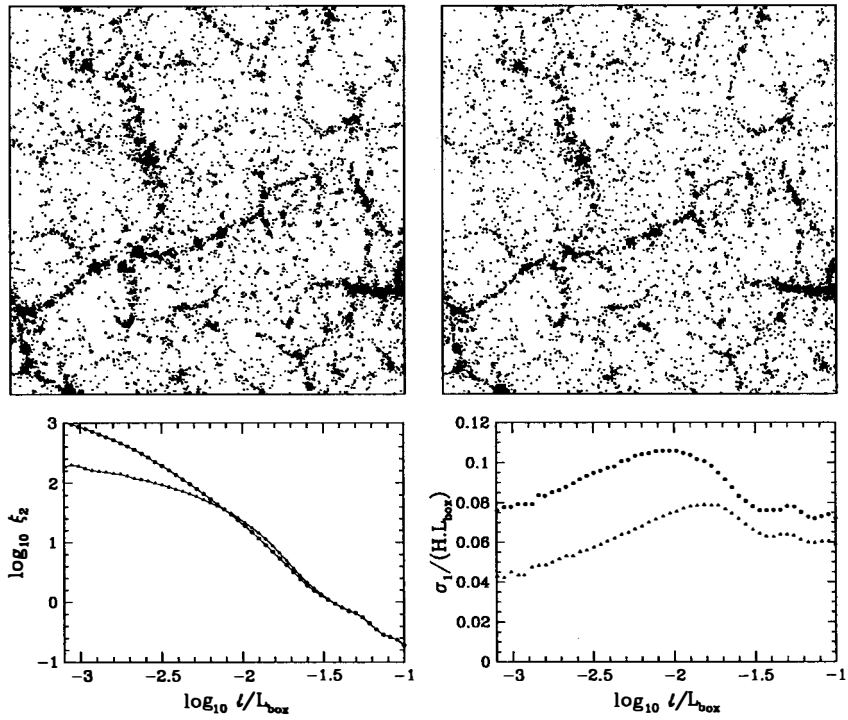


Figure 1: The two top panels represent thin slices of width $L_{\rm box}/16$ extracted from a PM simulation (left panel) and a tree simulation (right panel) with same initial conditions (see text). Bottom left and right panels compare respectively the measured function ξ_2 and the measured line-of-sight pairwise velocity dispersion (in units of the Hubble velocity across the box), in both simulations. Triangles and circles correspond respectively to PM and tree.

a function of scale in both simulations. Right bottom panel displays the line-of-sight pairwise velocity dispersion

$$\sigma_1(x) \equiv \frac{1}{\sqrt{3}} \langle [\vec{v}(\vec{y}+\vec{x}) - \vec{v}(\vec{y})]^2 \rangle^{1/2} \tag{18}$$

as a function of separation. As one can see, spatial resolution effects seem to have less impact on the mass distribution than on velocities, which is not very surprising. Note thus the good agreement of the measured two-point function in both simulations up to the PM resolution scale. There is however a $\sim$ 15 to 25% difference between the values of σ_1 measured in each simulations for $\ell \gtrsim \Delta x$, even at separations far above Δx. Of course, the quantity σ_1 is quite sensitive to the highly nonlinear details of the dynamics, since it is a mass weighted quantity much influenced by what happens in rich clusters of particles.

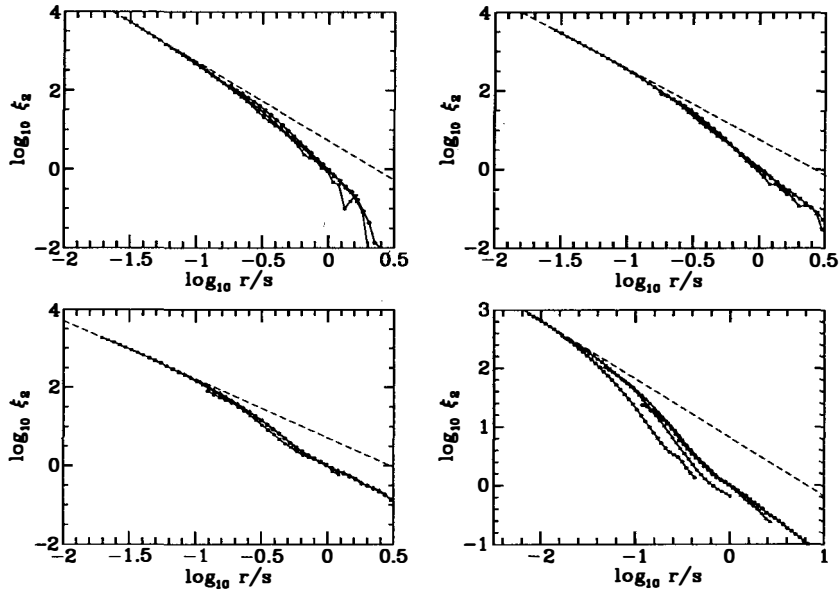


Figure 2: The quantity ξ_2 as function of the similarity variable, measured in scale free simulations[12]. The power-spectrum index $n = +1, 0, -1, -2$ decreases from left to right and from top to bottom. On each panel, various curves correspond to various stages of the simulations. The dashed line is the power-law (15) (the normalization is adjusted to the data).

5 Theoretical prejudices versus N-body experiments

One can compare the theoretical prejudices of § 3 to measurements in N-body simulations. The problem is to be aware of all the possible, unphysical numerical effects susceptible to contaminate the measurements. These spurious effects can be listed as follows:

- Dynamical effects (for example two-body relaxation) and lack of statistics at small scales due to the discreteness of the representation of the phase-space distribution function (the mass resolution cannot be arbitrarily large).
- Lack of accuracy at small scales due to the short range softening of the forces (the spatial resolution is finite).
- Memory from a particular way of setting up initial conditions. For example, there are *grid effects*[13] and transients resulting from the use of the Zel'dovich approximation to slightly perturb a regular pattern of particles to set up initial conditions. The use of a regular pattern is related to the fact one wants to reduce small scale shot noise. Grid effects can be eliminated by setting up the particles on a "glass"[14].
- Finite volume effects from the fact that only a finite part of the universe can be simulated.

Figure 2 shows the measured quantity ξ_2 as a function of r/t^α in scale-free simulations with initial power-spectra of the form (11). Figure 3 displays the measured ratios S_Q, $3 \leq Q \leq 5$, as

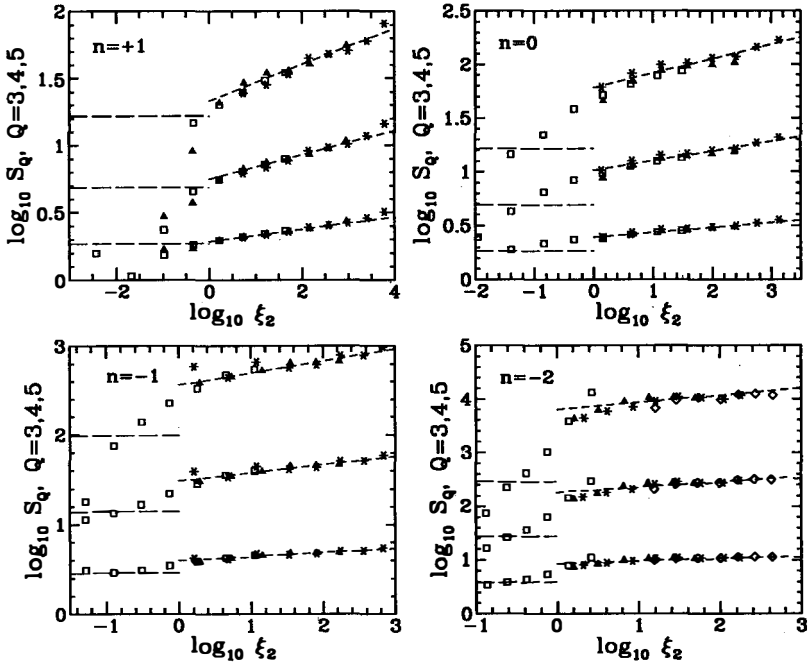


Figure 3: The ratios S_Q, $3 \leq Q \leq 5$, as functions of the variance (S_Q increases with Q), measured in scale-free simulations[12]. Various symbols correspond to various stages of the simulations. Long dashes correspond to perturbation theory predictions [except for $n = +1$, see ref. 12)]. Short dashes correspond to the following power-law fit $S_Q \propto \bar{\xi}_2^{0.045(Q-2)}$, in slight disagreement with the stable clustering hypothesis predictions S_Q =constant.

functions of variance $\bar{\xi}_2$. The simulations were done with a treecode[15] in a cubical box of size $L_{\rm box}$ with periodic boundaries. They involve $N_{\rm par} = 64^3$ particles and their spatial resolution is $\varepsilon = L_{\rm box}/1280$. To do the measurements, a "trustable" scaling range has been selected, where the contamination effects listed above are of small importance. In some cases, however, finite volume effects are so large that they affect all the available dynamic range. In fact, they are all the stronger since the ratio of large scale power to small scale power increases and since the order Q of the considered statistic is large. The consequence is that they need to be corrected for, generally as soon as $Q \geq 3$, as done in Fig. 3. Various symbols on each panel of Figs. 2, 3 correspond to various times in the simulations. Except in the case $n = -2$ for function $\bar{\xi}_2$, which is strongly contaminated by finite volume effects, the symbols superpose to each other, in agreement with the similarity solution (6). If the stable clustering hypothesis applied, function $\bar{\xi}_2(\ell, t)$ should also exhibit the power-law behavior (15) at small scales (short dashes on Fig. 2), which is indeed the case for $\bar{\xi}_2 \gtrsim 100$ [17], and the ratios S_Q should not depend on scale. This last property does not seem to be exactly verified here, except possibly for $n = -1$ and $n = -2$. More appropriate is the following power-law behavior $S_Q \propto \bar{\xi}_2^{0.045(Q-2)}$, a very weak deviation from stable clustering.

Qualitatively, given all the uncertainties on the measurements, the N-body results are in good agreement with theoretical prejudices in the highly nonlinear regime[17]. Quantitatively, there seem to be some small deviations from the stable clustering predictions.

Acknowledgements: figure 1 have been generated from measurements in a PM simulation done by E. Hivon with the code of Moutarde et al. [ref. 18)] and a treecode simulation done by L. Hernquist with the code of Hernquist et al. [ref. 15)]. Results of § 5 and figure 3 are extracted from a work done in collaboration with F.R. Bouchet and L. Hernquist [ref. 12)].

References

1) See, e.g., Peebles, P.J.E., 1980, The Large-Scale Structure of the Universe, Princeton University Press, Princeton
2) Doroshkevich, A.G., Ryabenkin, V.S., Shandarin, S.F., 1973, Astrofizica, 9, 257
3) Balian, R., Schaeffer, R., 1989, A&A, 255, 1
4) Szapudi, I., Szalay, A., 1993, ApJ, 408, 43
5) Davis, M., Peebles, P.J.E., 1977, ApJS, 34, 425
6) Davis, M., Groth, E.J., Peebles, P.J.E., 1977, ApJ, 212, L107
7) See, e.g., Hockney, R.W., Eastwood, J.W, 1981, Computer Simulation Using Particles, MacGraw-Hill, New York
8) See, e.g., Efstathiou, G., Davis, M., Frenk, C.S., White S.D.M., 1985, ApJS, 57, 241
9) See, e.g., Appel, A.W., 1985, SIAM, J. Sci. Stat. Comput., 6, 85; Jernigan, J.G., 1985, in IAU Symposium 113, Dynamics of Star Clusters, ed. J. Goodman and P. Hut, Reidel, Dordrecht
10) See, e.g., Barnes, J., Hut, P., 1986, Nature, 324, 446; Hernquist, L., 1987, ApJS, 64, 715; Bouchet, F.R., Hernquist, L., 1988, ApJS, 68, 521
11) Hivon, E., Bouchet, F.R., Colombi, S., Juszkiewicz, R., 1995, A&A, 298, 643
12) Colombi, S., Bouchet, F.R., Hernquist, L., 1996, ApJ, in press (astro-ph/9508142)
13) Colombi, S., Bouchet, F.R., Schaeffer, R., 1995, ApJS, 96, 401
14) See, e.g., Baugh, C.M., Gaztañaga, E., Efstathiou, G., 1995, MNRAS, 274, 1049
15) Namely the treecode of Hernquist, L., Bouchet, F.R., Suto, Y., 1991, ApJS, 75, 231
16) Colombi, S., Bouchet, F.R., Schaeffer, R., 1994, A&A, 281, 301
17) See also Efstathiou, G., Frenk, C.S., White, S.D.M., Davis, M., 1988, MNRAS, 235, 715
18) Moutarde, F., Alimi, J.-M., Bouchet, F.R., Pellat, R., Ramani, A., 1991, ApJ, 382, 377

Résumé

LE REGIME HAUTEMENT NON LINÉAIRE ET LES SIMULATIONS À N CORPS

Je discute de la dynamique à grande échelle dans un univers en expansion composé de matière non collisionnelle autogravitante, en prêtant particulièrement attention au régime hautement non lineaire, qui est généralement traité en utilisant des simulations à N corps. Mon approche est statistique, en termes de fonctions de corrélation et de fonction de distribution du champ de densité lissé. Après avoir rappelé les préjugés théoriques attendus (autosimilarité, équilibre statistique local), je discute de mesures dans des simulations à N corps.

HYDRODYNAMIC SIMULATIONS OF GALAXY FORMATION

Giuseppe Tormen

Institute of Astronomy, University of Cambridge - ENGLAND
and Max-Planck-Institut für Astrophysik, Garching - GERMANY

Abstract

This review is a short introduction to numerical hydrodynamics in a cosmological context, intended for the non specialist. The main processes relevant to galaxy formation are first presented. The fluid equations are then introduced, and their implementation in numerical codes by Eulerian grid based methods and by Smooth Particle Hydrodynamics is sketched. As an application, I finally show some results from an SPH simulation of a galaxy cluster.

1 Gas Physical Processes

In current cosmological scenarios, the main matter components of the universe are some non baryonic Dark Matter (DM), which constitutes most of the universe, and a mixture of primordial gas (H, He). The DM component is decoupled from the rest of the universe, and interacts only through gravity. The gas component instead can be heated and cooled in several ways, and the physics involved is more complicated than in the pure gravitational case. Fortunately, while gravity is a long range force, hydrodynamic processes are only important on relatively small scales, so that on scales larger than a few Mpc the dynamics of structure formation can be studied with good accuracy even neglecting the gas component. Gas dynamics, and the related radiative processes, are instead fundamental on smaller scales, e.g. in the formation of galaxies, and in linking the matter distribution of the universe to the light distribution we actually observe[1]. In what follows I list and discuss very briefly the main gas processes relevant to galaxy formation.

1.1 Heating processes

Adiabatic compression is the easiest way to heat a gas. By the First Law of Thermodynamics, compression work is converted into internal energy: $dQ = dU + pdV = 0 \Longrightarrow dU = -pdV$.

Viscous heating is due to the small internal friction (viscosity) present in real gas. Velocity gradients in a gas cause an irreversible transfer of momentum from high velocity points to small velocity ones, with conversion of bulk velocities into random ones, i.e. into heat, and generation of entropy. In the context of galaxy formation viscous heating mostly occurs in shocks, which are discontinuities in the macroscopic fluid variables due to supersonic flows. These arise for example during gravitational collapse, or during supernova (SN) explosions.

Photoionization takes place when atoms interact with the photons of some background of soft X-rays, or UV radiation emitted by QSO or stars: e.g. $\gamma + H \rightarrow e^- + H^+$. Observations show that at high redshift ($z \gtrsim 2$) hydrogen in the IGM is indeed ionized (Gunn-Peterson effect[2]), and although the origin is not clear, this is thought to be caused by some early generation of QSO or massive stars. The photoionizing spectrum is usually approximated by a power law, with flux $J(\nu) \propto (\nu/\nu_L)^{-\alpha}$, where ν_L is the Lyman-α frequency, corresponding to the hydrogen ionization energy, 13.6 eV.

1.2 Cooling processes

Gas cooling is the key to galaxy formation. In fact, in our current understanding of structure formation, the dark matter component of the universe first undergoes gravitational collapse, forming dark matter halos. These provide the potential wells into which gas can fall and heat up by shocks, then immediately cool and form cold, dense, rotationally supported gas disks. In these disks conditions are favourable to trigger star formation, and eventually give rise to the galaxies we observe today[3]. The following cooling mechanisms are important in a cosmological context.

Adiabatic expansion is the opposite process of adiabatic compression, with conversion of heat into expansion work.

Compton cooling is electron cooling against the Cosmic Microwave Background (CMB) through inverse Compton effect: $\gamma + e^- \rightarrow \gamma + e^-$. The condition for this is that the temperature of the electrons is higher than the CMB temperature: $T_e > T_\gamma$. The net energy transfer depends on the two densities, and on the temperature difference: $dE/dT \propto n_e \rho_\gamma (T_e - T_\gamma)$. Since $E \propto n_e T_e$, the cooling time: $t_{cool} \equiv E/\dot{E}$ is in this case $\propto \rho_\gamma^{-1}$. The CMB photon density decreases like $(1+z)^4$ due to the expansion of the universe, therefore Compton cooling is only important at high redshifts (typically $z \gtrsim 8$), when the cooling timescale is smaller than the Hubble time.

Radiative cooling is the most relevant mechanism for the cooling of primordial gas. It is caused by inelastic collisions between free electrons and H, He atoms (or their ions). Assuming that the gas is optically thin and in ionization equilibrium, the cooling rate per unit volume may be written $dE/dt \equiv \Lambda(\rho, T) = n_e n_i f(T)$, where n_e and n_i are the number densities of free electrons and of atoms (or ions), and $f(T)$ is called *cooling function*. The main processes are:

- *Collisional ionization:* the inelastic scattering of a free electron and an atom (or ion), which unbinds one electron from the latter, e.g. $e^- + H \rightarrow H^+ + 2e^-$. The net cooling for the system is equal to the extraction energy.
- *Collisional excitation + line cooling:* the same situation as above, but the atom is only excited, and it then decays to the ground state, emitting a photon, e.g. $e^- + H \rightarrow e^- + H^* \rightarrow e^- + H + \gamma$. This is the dominant cooling process at low (10^4 K $\lesssim T \lesssim 10^6$ K) temperatures.
- *Recombination:* e.g. $e^- + H^+ \rightarrow H + \gamma$. A free electron is captured by an ion and emits a (continuum) photon.
- *Bremsstrahlung:* free-free scattering between a free electron and an ion, e.g. $e^- + H^+ \rightarrow e^- + H^+ + \gamma$. Its cooling rate grows with the temperature: $dE/dt \propto T^{1/2}$; therefore, bremsstrahlung is the dominant cooling process at high ($T \gtrsim 10^6 K$) temperatures.

Figure 1 shows the cooling and heating functions in different cases.

1.3 Other processes

Besides the heating and cooling mechanisms listed above, other processes may be relevant in a galaxy formation scenario. Among them:

Thermal conduction is direct transfer of heat from regions at high temperature to regions at lower temperature, due to the energy transport of diffusing electrons. The induced change in internal energy per unit volume is $dE/dT = \vec{\nabla}\left(\kappa \vec{\nabla} T\right)$, with (positive) thermal conductivity $\kappa = \kappa(T, p)$.

Radiation transfer is important in an optically thick medium. Photons are absorbed by gas clouds, thermalized by multiple scattering and reemitted as Black Body radiation. This process depends on the optical depth τ_{opt} of the gas.

Star formation follows gas cooling as the next natural step in modeling galaxy formation. Our understanding of the detailed physics of star formation is still rather poor, so what is usually done is to use some empirical prescription to characterize the gas which is

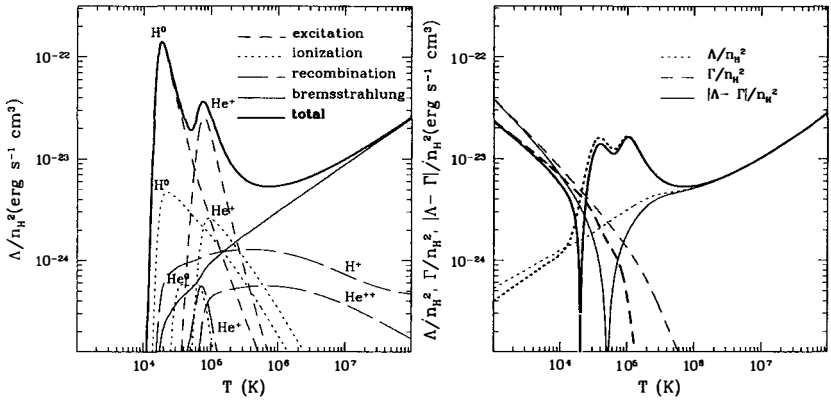


Figure 1: Cooling and heating functions. Left panel: cooling function $\Lambda(T)/n_H^2$ versus temperature, for a primordial gas. The different contributions to radiative cooling, and the total cooling curve are indicated. No photoionizing background is assumed. Right panel: a heating term Γ/n_H^2 is added, due to a photoionizing UV background with spectral index $\alpha = 1.5$. Its effect is both to change the ionization equilibrium (and so to change $\Lambda(T)/n_H^2$), and to heat the gas. Thin lines correspond to a gas density equal to the mean background density at redshift $z = 5$; thick lines correspond to a density 200 times larger. For low density gas the effect of a UV background is dramatic, and line cooling is suppressed. The temperature T_{eq} where $|\Lambda - \Gamma|$ drops to zero is called equilibrium temperature. At $T > T_{eq}$ the gas is cooled, at $T < T_{eq}$ it is heated. This figure was kindly prepared by I.Forcada using the atomic rates provided by T.Abel.

supposed to turn into stars. A good example of recipe[4)] is to require: *i)* a convergent gas flow: $\vec{\nabla} \cdot \vec{v} < 0$; *ii)* a Jeans' instability criterion: the free-fall time of the gas cloud be less than its sound crossing time; *iii)* a minimum number density of H atoms, e.g. $n_H > 0.1$ cm^{-3}; *iv)* a minimum gas overdensity, e.g. $\rho_g \gtrsim \rho_V = 178\bar{\rho}_g$, where ρ_V is the virial density of the spherical collapse model, and $\bar{\rho}_g$ is the mean background gas density. If all these conditions are satisfied, the gas will form stars at a formation rate similar to the one observed in e.g. spiral galaxies.

Assuming some Initial Mass Function for the stars so formed, it is possible to compute the fraction of gas forming massive stars ($M \geq 8M_\odot$). These stars will explode as Type II SN, each one releasing 10^{51} erg of energy back in the ISM, causing new shocks and gas heating, and triggering further star formation. It is also possible to include in this recipe gas release and metal enrichment from SN explosions. Unfortunately, at this point the physics of star formation is still poorly understood, and the resulting scenarios depend very much on the kind of assumptions and modeling made.

2 Fluid Equations

For our purpose it is useful to treat the gas as a continuum, and to resort to a hydrodynamic description. The fluid equations express conservation of mass (continuity equation), of momentum (Euler equation) and of energy. We also need an adiabatic state equation: $ds/dt = 0$ or

$p = p(\rho, T)$. These equations may be written as

$$\frac{d\rho}{dt} = -\rho\vec{\nabla}\cdot\vec{v}, \tag{1}$$

$$\rho\frac{d\vec{v}}{dt} = -\vec{\nabla}p + \text{viscosity terms} - \rho\vec{\nabla}\Phi, \tag{2}$$

$$\rho\frac{d\epsilon}{dt} = -p\vec{\nabla}\cdot\vec{v} + \text{viscosity terms} + \vec{\nabla}\cdot\left(\kappa\vec{\nabla}\mathrm{T}\right) + (\mathrm{Q} - \Lambda), \tag{3}$$

$$p = p(\rho, T). \tag{4}$$

Equation (3) is given in terms of the specific internal energy ϵ. In Equations (1) to (3) the *lhs* and the first term on the *rhs* constitute the usual fluid equations for a perfect adiabatic gas. The extra terms on the *rhs* are the nonadiabatic terms introduced in the previous Section. Gravity is included in the Euler equation via the gravitational potential Φ. Artificial viscosity terms must be included to enable the numerical treatment of shocks and entropy production, processes that do not exist in a perfect adiabatic gas. The terms Q and Λ denote respectively the heat sources: photon absorption (e.g. photoionization) and energy feedback from SN, and the heat sinks: Compton and radiative cooling, and other photon emission processes. Including all these processes (and others, e.g. magnetic fields) in a recipe for galaxy formation is not an easy task. Limits arise both from the computational limitations of present day machines, and from our ignorance of the physics involved (e.g. the epoch and spectrum of a photoionizing background, or the mechanism and role of energy feedback from stars and SN). Moreover, some processes naturally go with others, so that if cooling is implemented, star formation and energy feedback should also be modeled. For these reasons, different physical processes may or may not be taken into account in different computations. For example, current hydrodynamic simulations of structure formation sometimes include cooling processes, less often photoionization and star formation. On the other hand, thermal conduction and radiative transport have been so far neglected. The former may play a role in the central part of a system, but its efficiency would depend on the presence of (unknown) magnetic fields. Ignoring the latter is probably a good approximation everywhere except in very dense, optically thick regions. Galaxy cluster formation is a neat application of these methods, because cooling and star formation are less crucial here than in galaxy formation, so they can be ignored to first approximation. An example of hydrodynamic simulation of the formation of a galaxy cluster is briefly presented in the last Section of the paper.

2.1 Eulerian Methods

There are two basic ways to numerically solve the above set of equations: Eulerian and Lagrangian. Eulerian methods solve the fluid equations on a discrete grid fixed in space. In the traditional formulation, a Taylor expansion of the terms is used to build a smooth solution of the differential equations across different grid cells, as follows. Let f be a one dimensional scalar field, and F its flux; the conservation equation for F is

$$\frac{\partial f}{\partial t} = -\frac{\partial F}{\partial x}; \tag{5}$$

if we Taylor expand $f(x,t)$ in time:

$$f(x, t + dt) = f(x, t) + \frac{\partial f}{\partial t}dt + \frac{1}{2}\frac{\partial^2 f}{\partial t^2}dt^2 + O(dt^3) \tag{6}$$

and insert Eq. (5) into Eq. (6), we can substitute all time derivatives with spatial ones:

$$f(x,t+dt) = f(x,t) - \frac{\partial F}{\partial x}dt + \frac{1}{2}\frac{\partial}{\partial x}\left[\frac{\partial F}{\partial x}\frac{\partial F}{\partial f}\right]dt^2 + O(dt^3). \tag{7}$$

This is the time evolution equation for $f(x,t)$, which is discretized and solved on a grid, to first or second order accuracy. Macroscopic discontinuities in the flow, like those caused by shocks, are mimicked by smooth solutions, obtained introducing explicitly, in the fluid equations, the artificial viscosity terms mentioned above. The underlying idea is that in a real fluid shocks are smooth solutions if seen at small enough scale.

More recent techniques, named *shock capturing schemes* or *Riemann solvers*, use a different approach. They incorporate in the method the exact solution of a simple nonlinear problem, the Riemann shock tube. This solution describes the nonlinear waves generated by a discontinuous jump separating two constant states. The fluid flow is then approximated by a large number of constant states for which the Riemann shock tube problem is solved. This automatically leads to an accurate approximation of both smooth solutions and of shocks, with no need of explicitly introducing viscosity terms in the equations. This category of techniques includes the Piecewise Parabolic Method[5] (PPM) and the Total Variation Diminishing[6] scheme (TVD).

2.2 Lagrangian methods: SPH

Smooth Particle Hydrodynamics[7),8)] (SPH) is the most commonly used Lagrangian method in dissipative simulations of galaxy formation. By analogy with N-body codes, where a collisionless fluid is represented with a set of discrete particles, SPH also uses particles to describe the evolution of a gas fluid. Each particle i is assigned a position $\vec{r}_i$, a velocity $\vec{v}_i$, a density ρ_i and a specific internal energy ϵ_i, and the fluid equations are solved at the particle's position, replacing the true fields with smoothed estimates, which are obtained as local averages of the particles' properties. For example, for a scalar field f, the smoothed counterpart $\langle f \rangle$ is:

$$\langle f(\vec{r})\rangle = \int d^3u f(\vec{u}) W(\vec{r}-\vec{u};h); \tag{8}$$

$W(\vec{r}-\vec{u};h)$ is the *smoothing kernel*, strongly peaked at zero, so that

$$\lim_{h\to 0}\langle f(\vec{r})\rangle = \int d^3u f(\vec{u})\delta_D(\vec{r}-\vec{u}) = f(\vec{r}), \tag{9}$$

with δ_D a 3 dimensional Dirac delta; h is called the *smoothing length*. Usually the kernel is spherically symmetric: $W = W(|\vec{r}-\vec{u}|;h)$, but anisotropic kernels have also been proposed[9)].

In practice, $\langle f(\vec{r})\rangle$ is evaluated discretely at each particle's position. Defining the particle number density at $\vec{r}_i$ as $\langle n(\vec{r}_i)\rangle = \rho(\vec{r}_i)/m_i$, in the discrete limit Equation (8) becomes

$$\langle f(\vec{r})\rangle = \sum_{j=1}^{N} \frac{m_j}{\rho_j} f(\vec{r}_j) W(|\vec{r}-\vec{r}_j|;h); \tag{10}$$

for example, for $f=\rho$ this reduces to

$$\langle \rho(\vec{r})\rangle = \sum_{j=1}^{N} m_j W(|\vec{r}-\vec{r}_j|;h). \tag{11}$$

Typically, in SPH codes the kernel is nonzero only for e.g. $|\vec{r}-\vec{u}| \le 2h$, and the smoothing length can be varied to keep the summations to the $N = 30-50$ closest neighbors.

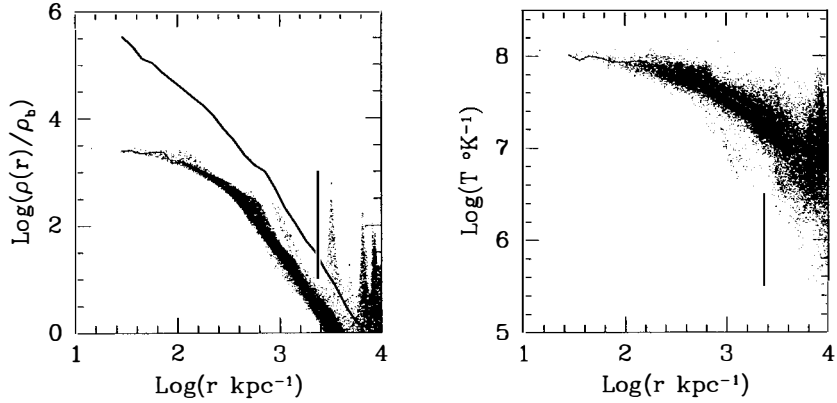


Figure 2: Density (left) and temperature (right) radial profiles for an SPH simulation of a galaxy cluster. The cluster has a virial mass of $7 \times 10^{14} M_{\odot}$ and is resolved by roughly 12000 particles for each species. The force resolution is 25 kpc. More details on the figure are given in the main text.

Following this recipe, one can write the fluid equations: mass conservation is automatically satisfied; one form of the Euler and energy equations is, in the adiabatic case[7]:

$$\frac{d\vec{v}_i}{dt} = -\sum_{j=1}^{N} m_j \left[\frac{p_i}{\rho_i^2} + \frac{p_j}{\rho_j^2}\right] \vec{\nabla}_i W(|\vec{r}_i - \vec{r}_j|; h); \tag{12}$$

$$\frac{d\epsilon_i}{dt} = \frac{p_i}{\rho_i^2} \sum_{j=1}^{N} m_j (\vec{v}_i - \vec{v}_j) \cdot \vec{\nabla}_i W(|\vec{r}_i - \vec{r}_j|; h). \tag{13}$$

Nonadiabatic terms are introduced in the same way; in particular, shock heating is allowed by adding artificial viscosity terms.

Hybrid methods have also been proposed[10], that make use of a grid to solve the fluid equations, but are Lagrangian in nature, since the grid itself is deformable and follows the fluid flow. The quite different approaches of Eulerian and Lagrangian methods, and the variety of codes within each approach, make it difficult to compare results coming from different codes, and to interpret the comparison. An attempt has however been done[11] by evolving, from some high redshift to the present time, the same initial conditions of a cosmological model, using three different Eulerian codes and two different SPH codes; the results seem to indicate that Eulerian codes have better resolution on large scales and low ρ, low T regions, while SPH codes can better resolve small scales and high ρ, high T regions.

3 Applications: Formation of a galaxy cluster

Some results from an SPH simulation of a galaxy cluster are shown in Figure 2, as an example of application of the theory summarized in the previous Sections. The gas processes modeled are adiabatic heating and cooling, and viscous heating, while other cooling processes, photoionization and star formation were neglected, since they are not crucial to this particular problem.

The simulation has an Einstein-de Sitter background universe, with $H_0 = 50\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$, and with scale free power spectrum of perturbation $P(k) \propto k^{-1}$. The left panel shows the gas density (points), and the DM and gas mean density profiles (black and grey curve respectively), as a function of the distance r from the cluster center; the vertical bar indicates the virial radius of the cluster. Note how the gas is less centrally concentrated than the DM. The density spikes in the gas profile are infalling substructure. The right panel shows the gas temperature and the mean temperature profile at different radii: the cluster is not isothermal, and its temperature drops by a factor of five from the center to the virial radius. During collapse the gas is shock heated to about 10^7 - 10^8 K as seen both in the main system and in the infalling substructures.

Acknowledgments

I would like to thank Bruno Guiderdoni for inviting me to give this talk. Thanks also to Bhuvnesh Jain, Ignasi Forcada and Simon White for comments on the manuscript. Financial support from an EC-HCM fellowship is gratefully acknowledged.

References

[1] White, S.D.M., 1995, in "Lecture Series of the Les Houches Summer School on 'Dark Matter and Cosmology" August 1993, ed. R. Schaeffer, North Holland, in press.

[2] Gunn, J.E. and Peterson, B.A., 1965, ApJ 142, 1633.

[3] Rees, M.J. and Ostriker, J.P., 1977, MNRAS 179, 541.

[4] Katz, N., Weinberg, D.H. and Hernquist, L., 1995, ApJS submitted.

[5] Bryan, G. L., Norman, M. L., Stone, J. M., Cen, R., and Ostriker, J. P., 1995, Comp.Phys.Comm. 89, 149.

[6] Ryu, D., Ostriker, J.P., Kang, H. and Cen, R., 1993, ApJ 414, 1.

[7] Benz, W., 1990, in Numerical Modelling of Nonlinear Stellar Pulsations, ed. J.R. Buchler, Kluwer, Dordrecht.

[8] Steinmetz, M., 1996, in Proceedings "International School of Physics Enrico Fermi", Course CXXXII: Dark Matter in the Universe, Varenna 1995, IOP, in press.

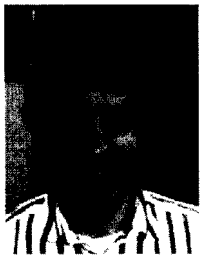
[9] Owen, M.O., Villumsen, J., Shapiro, P.R. and Martel, H., 1996, MNRAS sumbitted.

[10] Gnedin, N.Y., 1995, ApJS 97, 231.

[11] Kang, H. et al., 1994, ApJ 430, 83.

THE SLOAN DIGITAL SKY SURVEY: STATUS AND PROSPECTS

Jon Loveday
Fermilab, Batavia, USA.
(On behalf of the SDSS collaboration.)



Abstract

The Sloan Digital Sky Survey (SDSS) is a project to definitively map π steradians of the local Universe. An array of CCD detectors used in drift-scan mode will digitally image the sky in five passbands to a limiting magnitude of $r' \sim 23$. Selected from the imaging survey, 10^6 galaxies and 10^5 quasars will be observed spectroscopically. I describe the current status of the survey, which is due to begin observations early in 1997, and its prospects for constraining models for dark matter in the Universe.

1 Introduction

Systematic surveys of the local Universe ($z \lesssim 0.2$) can provide some of the most important constraints on dark matter, particularly through the measurement of the clustering of galaxies and clusters of galaxies on large scales. Most existing galaxy and cluster catalogues are based on photographic plates [7, 3], and there is growing concern that such surveys might suffer from severe surface-brightness selection effects, so that they are missing a substantial fraction of the galaxy population. In addition, the limited volume of existing redshift surveys means that even low-order clustering statistics, such as the galaxy two-point correlation function, cannot reliably be measured on scales beyond $100h^{-1}$Mpc, an order of magnitude below the scale on which COBE has measured fluctuations in the microwave background radiation.

A collaboration has therefore been formed with the aim of constructing a definitive map of the local universe, incorporating digital CCD imaging over a large area in several passbands and redshifts for around one million galaxies. In order to complete such an ambitious project over a reasonable timescale, it was decided to build a *dedicated* 2.5-metre telescope equipped

with a large CCD array imaging camera and multi-fibre spectrographs. The collaboration comprises around 100 astronomers and engineers from University of Chicago, Fermilab, Princeton University, Institute for Advanced Study, Johns Hopkins University, US Naval Observatory, University of Washington and the JPG—a group of astronomers in Japan. The total cost of the survey is around $30 million, and funding sources include the Alfred P. Sloan Foundation, the National Science Foundation and the participating institutions.

2 Survey Overview

The survey site is Apache Point Observatory, New Mexico, at 2800 metres elevation. While better sites probably exist in Chile and atop Mauna Kea, for a survey with such state-of-the-art instrumentation and significant on-site manpower requirements (eg. fibre plugging and changing spectroscopic plates), it was decided to use a site within mainland USA and with good communications and existing infrastructure.

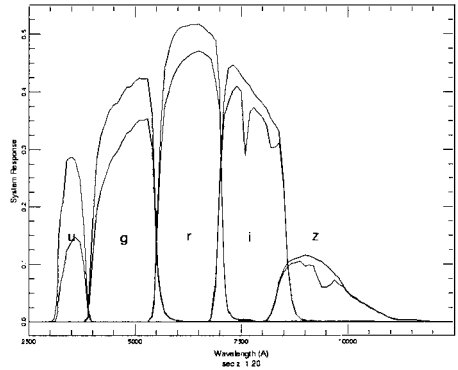


Figure 1: SDSS system response curves, with (lower) and without (upper) atmospheric extinction.

The survey hardware comprises the main 2.5-metre telescope, equipped with CCD imaging camera and multi-fibre spectrographs, a 0.6-metre monitor telescope and a 10μ all-sky camera. On the best nights (new moon, photometric, sub-arcsecond seeing) the 2.5-metre telescope will operate in imaging mode, drift scanning the sky at sidereal rate, and obtaining nearly simultaneously images in the five survey bands u', g', r', i' and z'. The system response curves through the five filters are shown in Figure 1. On sub-optimal nights, which will comprise the bulk of observing time, the imaging camera will be replaced with a spectroscopic fibre plug-plate. It is planned that imaging data will be reduced and calibrated, spectroscopic targets selected, and plates drilled within the one-month lunar cycle, so that we will be obtaining spectra of objects that were imaged the previous month. We will spend most of the time observing within a contiguous π steradian area in the north Galactic cap (NGC). For those times when the NGC is unavailable, about one third of the time, we will repeatedly observe three southern stripes, nominally centred at RA $\alpha = 5°$, and with central declinations of $\delta = +15°$, $0°$ and $-10°$. The nominal location of survey scans is shown in Figure 2.

In the remainder of this section I discuss the various components of the survey in more detail.

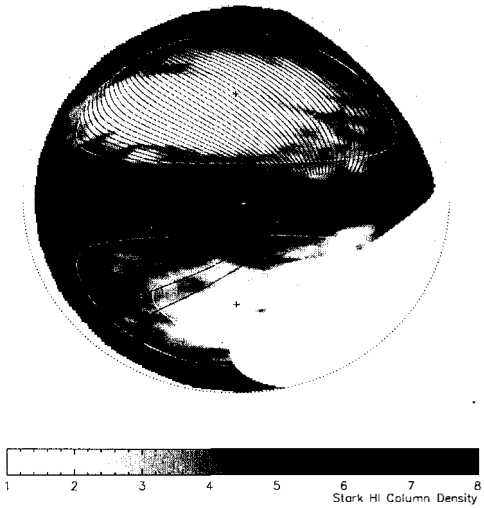


Figure 2: Whole sky plot showing the location of SDSS scans. The light lines show galactic latitudes of $b = 0$, ± 30 and $\pm 60°$, with the north and south Galactic poles being the upper and lower crosses respectively. The Galactic plane runs horizontally through the middle of the plot and the grey scale map shows Stark HI contours in units of 10^{20} cm^{-2}. The dark lines show the survey scan-lines, all of which follow great circles. We observe a contiguous area of π sr in the north, and three separated stripes in the south. Note that the northern survey is tilted with respect to the $b = +30°$ contour to avoid regions of high HI column density.

2.5-metre telescope. The main 2.5-metre telescope is of modified Richey-Chretien design with a 3° field of view, and is optimised for both a wide-area imaging survey and a multi-fibre spectroscopic survey of galaxies to $r' \sim 18$. One of the most unusual aspects of the telescope is it's enclosure. Rather than sitting inside a dome, as is the case with conventional optical telescopes, the enclosure is a rectangular frame structure mounted on wheels, which is rolled away from the telescope in order to take observations. By completely removing the enclosure from the telescope, we can avoid the substantial degradation to image quality due to dome seeing. The telescope is situated on a pier overlooking a steep dropoff so that the prevailing wind will flow smoothly over the telescope in a laminar flow, which will also help to ensure good image quality. A wind baffle closely surrounds the telescope, and is independently mounted and driven. This baffle serves to protect the telscope from stray light as well as from wind buffeting.

Imaging Camera. In order to image a large area of sky in a short time, we are building an imaging camera (Fig. 3) that contains 30×2048^2 CCDs, arranged in six columns. Each column occupies its own dewar and contains one chip in each of the five filters. Pixel size is 0.4″. The camera operates in drift-scan mode: a star or galaxy image drifts down the column through the five filters, spending about 55 seconds in each. This mode of observing has two significant advantages over conventional tracking mode. 1) It makes extremely efficient use of observing time, since there is no overhead between exposures: on a good night we can open the shutter, drift-scan for eight hours and then close the shutter. 2) Since each image traverses a whole column of pixels on each CCD, flat-fielding becomes a one-dimensional problem, and so can be done to lower surface-brightness limits than with tracking mode images. This, along with the high quantum efficiency of modern CCDs, will enable us to detect galaxies of much lower surface brightness than can wide-field photographic surveys. There is a gap between each column of CCDs, but this gap is slightly smaller than the width of the light-sensitive area of the

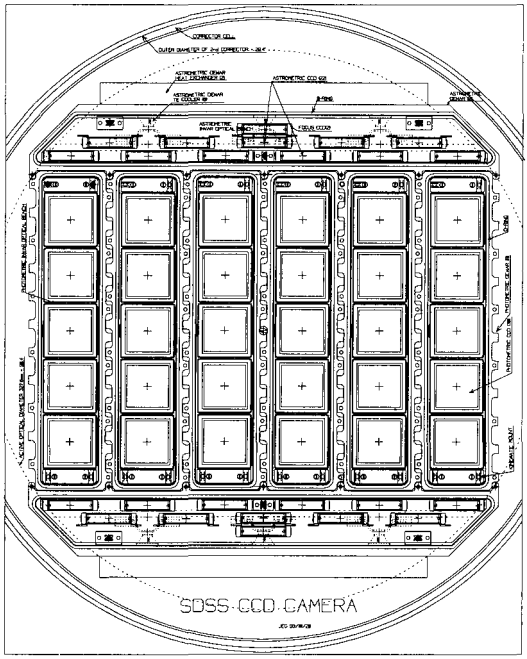


Figure 3: Focal plane layout of the SDSS CCD imaging camera, showing the 30 photometric and 24 astrometric/focus CCDs.

CCDs, and so having observed six narrow strips of sky one night, we can observe an interleaving set of strips a later night, and thus build up a large contiguous area of sky. The northern survey comprises 45 pairs of interleaving great circle scans, and so imaging observations for the north will require the equivalent of 90 full photometric nights. The camera also includes 24 smaller CCDs arranged above and below the photometric columns. These extra CCDs, equipped with neutral density filters, are used for astrometric calibration, as most astrometric standards will saturate on the photometric CCDs. Thus the photometric data can be tied to the fundamental astrometric reference frames defined by bright stars.

Spectrographs. The 2.5-metre telescope will also be equipped with a pair of fibre-fed, dual-beam spectrographs, each with two cameras, two gratings and two 2048^2 CCD detectors. The blue channel will cover the wavelength range 3900–6100 Å and the red channel 5900–9100 Å and both will have a spectral resolving power $\lambda/\Delta\lambda \approx 1800$. The fibres are 3″ in diameter and the two spectrographs each hold 320 fibres. Rather than employing robotic fibre positioners to place the fibres in the focal plane, we will instead drill aluminium plates for each spectroscopic field and plug the fibres by hand. We plan on spectroscopic exposure times of 45 minutes and allow 15 minutes overhead per fibre plate. On a clear winter's night we can thus obtain 8 plates $\times$ 640 fibres = 5120 spectra. In order to allow such rapid turnaround time between exposures we plan to purchase 8 sets of fibre harnesses, so that each plate can be plugged with fibres during the day. It will not be necessary to plug each fibre in any particular hole, as a fibre mapping system has been built which will automatically map fibre number onto position in the focal plane after the plate has been plugged. This should considerably ease the job of the fibre pluggers, and we expect that it will take well under one hour to plug each plate.

Monitor telescope. In order to check that observing conditions are photometric, and to

tie imaging observations to a set of primary photometric standards, we are also employing a monitor telescope. While the 2.5-metre telescope is drift-scanning the sky, the 0.6-metre monitor telescope, situated close by, will interleave observations of standard stars with calibration patches in the area of sky being scanned. Operation of this telescope will be completely automated, and each hour will observe three calibration patches plus standard stars in all five colours.

10μ **all-sky camera.** As an additional check on observing conditions, a 10μ infrared camera will survey the entire sky every 10 minutes or so. Light cirrus, which is very hard to see on a dark night, is bright at 10μ, and so this camera will provide rapid warning of increasing cloud cover, thus enabling us to switch to spectroscopic observing rather than taking non-photometric imaging data.

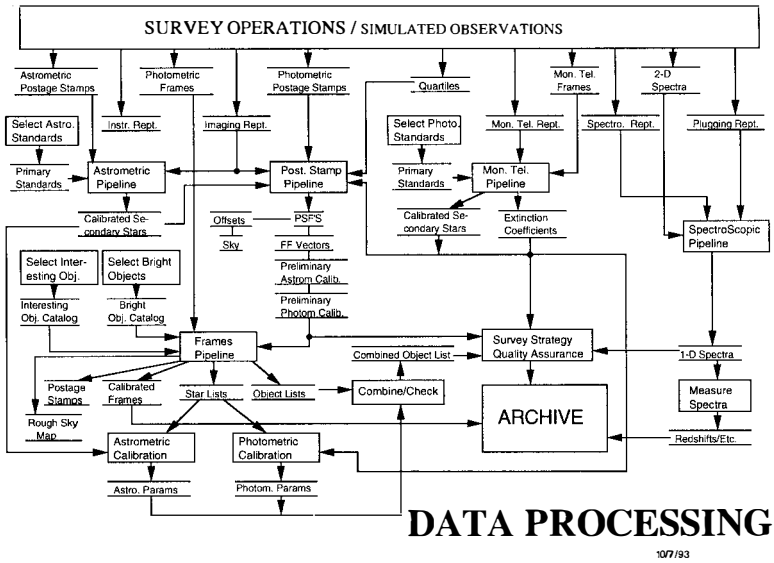


Figure 4: Top-level data processing diagram.

Data-reduction pipelines. The last, but by no means least, component of the survey is a suite of automated data-reduction pipelines (Fig. 4), which will read DLT tapes mailed to Fermilab from the mountain and yield reduced and calibrated data with the minimum of human intervention. Such software is very necessary when one considers that the imaging camera will produce data at the rate of around 31 Gbytes per hour! A "production system" has been specced and purchased that can keep up with such a data rate (bearing in mind that imaging will take place only under the best conditions, on average around two full nights per month), and consists of two Digital Alphaserver 8200 5/300s, each with 1 GByte of memory.

Pipelines exist to reduce each source of data from the mountain (photometric frames and "postage stamps", astrometric frames, monitor telescope frames and 2-D spectra) as well as to perform tasks such as spectroscopic target selection and "adaptive tiling" to work out the optimal placing of spectroscopic field centres to maximize the number of spectra obtained. The

pipelines are integrated into a purpose-written environment known as SHIVA (Survey Human Interface and VisualizAtion environment, also the Hindu god of destruction) and the reduced data will be written into an object-oriented database.

3 Data Products

The raw imaging data in five colours for the π steradians of the northern sky will occupy about 14 Tbytes, but it is expected that very few projects will need to access the raw data, which will probably be stored only on magnetic tape. Since most of the sky is blank to $r' \sim 23$, all detected images can be stored, using suitable compression, in around 200 Gbytes, and it is expected that these "atlas images" can be kept on spinning disc. The photometric reduction pipeline will meaure a set of parameters for each image, and it is estimated that the parameter lists for all objects will occupy $\sim$ 100 Gbyte. The parameter lists for the spectroscopic sample will proabbly fit into 1–2 Gb, and the spectra themselves will occupy $\sim$ 20 Gb. Work is progressing well on an astronomer-friendly interface to the database, which will answer such queries as "Return all galaxies with $(g' - r') < 0.5$ and within 30 arcminutes of this quasar", etc.

3.1 Spectroscopic Samples

The spectroscopic sample is divided into several classes. In a survey of this magnitude, it is important that the selection criteria for each class remain fixed throughout the duration of the survey. Therefore, we will spend a considerable time (maybe one year), obtaining test data with the survey instruments and refining the spectroscopic selection criteria in light of our test data. Then, once the survey proper has commenced, these criteria will be "frozen in" for the duration of the survey. The numbers discussed below are therefore only preliminary, and we expect them to change slightly during the test year.

The **main galaxy sample** will consist of $\sim 900,000$ galaxies selected by Petrosian magnitude in the r' band, $r' \lesssim 18$. Simulations have shown that the Petrosian magnitude, which is based on an aperture defined by the ratio of light within an annulus to total light inside that radius, provides probably the least biased and most stable estimate of total magnitude. There will also be a surface-brightness limit, so that we do not waste fibres on galaxies of too low surface brightness to give a reasonable spectrum. This galaxy sample will have a median redshift $\langle z \rangle \approx 0.1$.

We plan to observe an additional $\sim 100,000$ **luminous red galaxies** to $r' \lesssim 19.5$. Given photometry in the five survey bands, redshifts can be estimated for the reddest galaxies to $\Delta z \approx 0.02$ or better [4], and so one can also predict their luminosity quite accurately. Selecting luminous red galaxies, many of which will be cD galaxies in cluster cores, provides a valuable supplement to the main galaxy sample since 1) they will have distinctive spectral features, allowing a redshift to be measured up to 1.5 mag fainter than the main sample, and 2) they will form an approximately volume-limited sample with a median redshift $\langle z \rangle \approx 0.5$. They will thus provide an extremely powerful sample for studying clustering on the largest scales and the evolution of galaxies.

Quasar candidates will be selected by making cuts in multi-colour space and from the FIRST radio catalogue [1], with the aim of observing $\sim 100,000$ quasars. This sample will be orders of magnitude larger than any existing quasar catalogues, and will be invaluable for quasar luminosity function, evolution and clustering studies as well as providing sources for followup absorption-line observations.

In addition to the above three classes of spectroscopic sources, which are designed to provide *statistically complete* samples, we will also obtain spectra for many thousands of **stars** and for various **serendipitous** objects. The latter class will include objects of unusual colour or morphology which do not fit into the earlier classes, plus unusual objects found by other surveys and in other wavebands.

4 Current Status

In this section I discuss the status (as of April 1996) of the various systems within the survey.

The **monitor telescope** has been operational now for several months, and is routinely operated remotely from Chicago. It is equipped with a set of SDSS filters, and is being used to observe candidate primary photometric standard stars, as well as known quasars to see where they lie in the SDSS colour system [8].

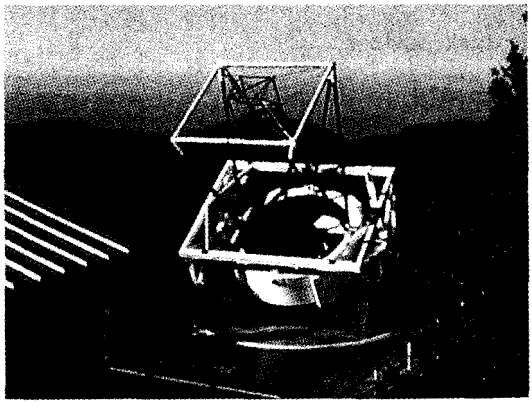


Figure 5: Photograph of the 2.5-metre telescope structure, taken shortly after installation, on 10 October 1995. Part of the telescope enclosure, in its rolled back position, appears in the bottom-left of this picture. Note that neither the mirrors nor the wind baffle are installed yet.

The **2.5-metre telescope** structure was installed on the mountaintop in October 1995 (see Fig. 5). Work is currently underway on the control systems for the telescope. Telescope **optics** are all due to be ready by June 1996. These include the primary and secondary mirrors and various corrector elements.

We posess all of the CCDs for the **imaging camera**, which is under construction at Princeton. Delivery to the mountain is expected by September 1996. Construction of the **spectrographs** is well underway, with the optics installed for one of the spectrograph cameras.

Each of the data reduction-pipelines is now basically working, with ongoing work on minor bug-fixes, speed-ups and integration of the entire data processing system. The photometric reduction pipeline is being tested using both simulated data and with data taken using the Fermilab drift scan camera on the ARC 3.5-metre telescope at the same site. Similar tests are being carried out on the spectroscopic reduction pipeline, and our ability to efficiently place fibres on a clustered distribution of galaxies is being tested using the APM galaxy catalogue [7].

The currently-projected survey schedule is as follows:

September 1996	Optics to be installed on 2.5-metre telescope.
Autumn 1996	Imager and spectrograph commissioning.
Winter 1996	Astronomical first light.
Early 1997	Test period begins.
1998–2003	Survey proper carried out.
2002	First two years of survey data become public.
2005	Complete survey data become public.

The intent of this project is to make the survey data available to the astronomical community in a timely fashion. We currently plan to distribute the data from the first two years of the survey no later than two years after it is taken, and the full survey no later than two years after it is finished. The first partial release may or may not be in its final form, depending on our ability to calibrate it fully at the time of the release. The same remarks apply to the release of the full data set, but we expect the calibration effort to be finished before that release.

5 Prospects for constraining dark matter

Since one of the topics of this meeting is dark matter, I will highlight two of the areas in which the SDSS will provide valuable data for constraining dark matter.

5.1 Measurement of the Fluctuation Spectrum

The huge volume of the SDSS redshift survey will enable estimates of the galaxy power spectrum to $\sim 1000h^{-1}$Mpc scales. Figure 6 shows the power spectrum $P(k)$ we would expect to measure from a volume-limited (to M^*) sample of galaxies from the SDSS northern redshift survey, assuming Gaussian fluctuations and a $\Omega h = 0.3$ CDM model. The error bars include cosmic variance and shot noise, but not systematic errors, due, for example, to galactic obscuration. Provided such errors can be corrected for, (and star colours in the Sloan survey will provide our best *a posteriori* estimate of galactic obscuration), then the figure shows that we can easily distinguish between $\Omega h = 0.2$ and $\Omega h = 0.3$ models, just using the northern main galaxy sample. Adding the southern stripe data, and the luminous red galaxy sample, will further decrease measurement errors on the largest scales, and so we also expect to be able to easily distinguish between low-density CDM and MDM models, and models with differing indices n for the shape of the primordial fluctuation spectrum.

5.2 Cosmological Density Parameter

By measuring the distortions introduced by streaming motions into redshift-space measures of galaxy clustering, one can constrain the parameter $\beta = \Omega^{0.6}/b$, where Ω is the cosmological density paramter and b is the bias factor relating fluctuations in galaxy number density to fluctuations in the underlying mass distribution. While existing redshift surveys, eg. IRAS [2] and Stromlo-APM [6], are hinting that $\beta < 1$ (ie. that galaxies are significantly biased tracers of mass or that $\Omega < 1$), their volumes are too small to measure galaxy clustering in the fully linear regime reliably enough to measure β to much better than 50% or so. With the SDSS redshift survey, we expect to be able to constrain β to 10% or better.

There are several ways we might hope to determine the galaxy bias factor b. By measuring galaxy clustering on $\sim 1000h^{-1}$Mpc scales as shown in Figure 6, we can compare with the COBE microwave background fluctuations directly, and so constrain large-scale galaxy bias in

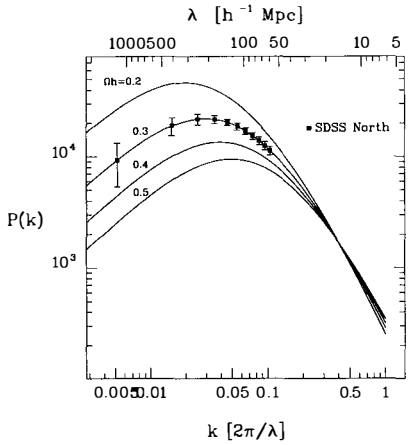


Figure 6: Expected 1σ uncertainty in the galaxy power spectrum measured from a volume-limited sample from the SDSS northern survey, along with predictions of $P(k)$ from four variants of the low-density CDM model. Note that the models have been arbitrarily normalised to agree on small scales ($k = 0.4$); in practice the COBE observations of CMB fluctuations fix the amplitude of $P(k)$ on very large scales.

a model-independent way. Analysis of higher-order clustering statistics [5], and of non-linear dynamical effects [2] will also set constraints on galaxy bias. Knowing β and b. we will be in a good position to reliably measure the cosmological density parameter Ω independent of models for the shape of the fluctuation spectrum.

6 Conclusions

It is probably no exaggeration to claim that the Sloan Digital Sky Survey will revolutionize the field of large scale structure. Certainly we can expect to rule out large numbers of presently viable cosmological models, as illustrated in Figure 6. As well as measuring redshifts for a carefully controlled sample of 10^6 galaxies and 10^5 quasars, the survey will also provide high quality imaging data for about 100 times as many extragalactic objects, from which one can obtain colour and morphological information. In addition to measuring the basic cosmological parameters Ω and h discussed in the preceding section, the SDSS will also allow us to measure the properties of galaxies as a function of their colour, morphology and environment, providing valuable clues to the process of galaxy formation.

Finally, I cannot resist the temptation to give a visual impression of what we might expect to see with the SDSS redshift survey. Figure 7 shows the distribution of 62,295 galaxies in a 6° slice from a simulation carried out by Changbom Park, assuming a low-density CDM model. This slice represents just *one sixteenth* of the million galaxy redshifts we will be measuring with the Sloan survey. I leave it to the readers imagination to dream up all the projects they would love to carry out given such a data-set.

The work described here has been carried out by many people throughout the SDSS collaboration, and I thank all my colleagues warmly. I am particularly grateful to Chris Stoughton and Michael Vogeley for providing Figures 2 and Figure 6 respectively, and to Philippe Canal for translating the Abstract into French. My attendance at the meeting was supported by a generous grant from the EEC.

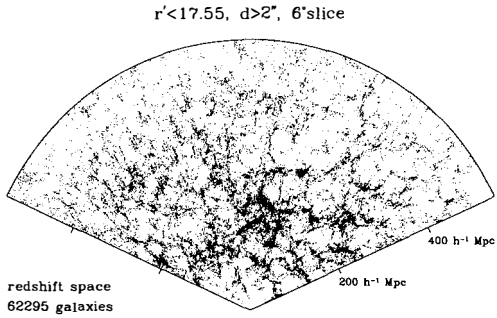


Figure 7: Redshift-space distribution of galaxies in a 6° slice from a large, low-density CDM N-body simulation generated by Changbom Park.

References

[1] Becker, R.H., White, R.L. and Helfand, D.J., 1995, ApJ, 450, 559

[2] Cole, S., Fisher, K.B. and Weinberg, D.H., 1995, MNRAS, 275, 515

[3] Collins, C.A., Heydon-Dumbleton, N.H. and MacGillivray, H.T., 1989, MNRAS, 236, 7P

[4] Connolly, A.J., et al., 1995, AJ, 110, 2655

[5] Gaztañaga, E. and Frieman, J.A., 1994, ApJ, 437, L13

[6] Loveday, J., Efstathiou, G., Maddox, S.J. and Peterson, B.A., 1996, ApJ, in press

[7] Maddox, S.J., Sutherland, W.J. Efstathiou, G. and Loveday, J., 1990, MNRAS, 243, 692

[8] Richards, G.T., et al., PASP, submitted

LE SLOAN DIGITAL SKY SURVEY: L'ÉTAT ET CES PERSPECTIVES

Le Sloan Digital Sky Survey (SDSS) à pour but de cartographié π steradians de l'univers local. Une matrice de dispositif à transfert de charges (CCD) scannant en mode balayage produira une image digitalisée du ciel avec cinq différents filtres et avec une précision allant jusqu'a à peu près magnitude 23. Une étude spectroscopique sera faite sur une sélection de 10^6 galaxies et 10^5 quasars. Dans cet article, après avoir décris l'état d'advancement du projet qui doit commencer à faire des observations des le début de l'année 1997, je présente ces perspectives pour l'établissement de modèles de la matière noire dans l'univers.

The DENIS & 2MASS Near Infrared Surveys and their Applications in Cosmology

GARY MAMON
IAP, Paris & DAEC, Obs. de Paris, Meudon

Abstract

The DENIS and 2MASS near infrared surveys are presented. Their applications in extragalactic astronomy and cosmology are listed. The prospects for a rapid spectroscopic followup survey of a near infrared selected sample of nearly 10^5 galaxies are illustrated with Monte-Carlo simulations.

1. Introduction

Astronomers have traditionally relied upon optical imagery to view the Universe, ever since the early introduction of photographic plates, sensitive to blue light. The main constituents of the Universe in the optical are *galaxies*, and the main constituents of galaxies in the optical are *stars*. In fact, the visible parts of galaxies contain important amounts of gas and non-negligible amounts of dust, and perhaps some dark matter. Moreover, most of the mass in galaxies is thought to be constituted of dark matter residing in near spherical halos. Similarly, on the scale of the Universe, most of the baryons are presumed to be locked up in gas, and most of the mass is believed to be non-baryonic dark matter. Working in the optical, one uses galaxies to trace the matter distribution in the Universe. But galaxies may very well be biased tracers of the underlying matter content of the Universe.

Near infrared (NIR) light is also sensitive to stars, hence to galaxies. But there are important differences.

1) NIR light is typically 10 times less extinguished by dust than optical light. This means that the cores of galaxies, hidden in the optical, are visible in the NIR. This is illustrated in Figure 1, which shows the same galaxy, in blue (B) and NIR (K') light. The white pattern cutting through the galaxy in the blue image, is caused by internal extinction by dust, and is invisible in the NIR image. The galaxy is much more symmetric in NIR light and the shapes of the isophotes from disky inside to boxy outside can only be seen on the NIR image.

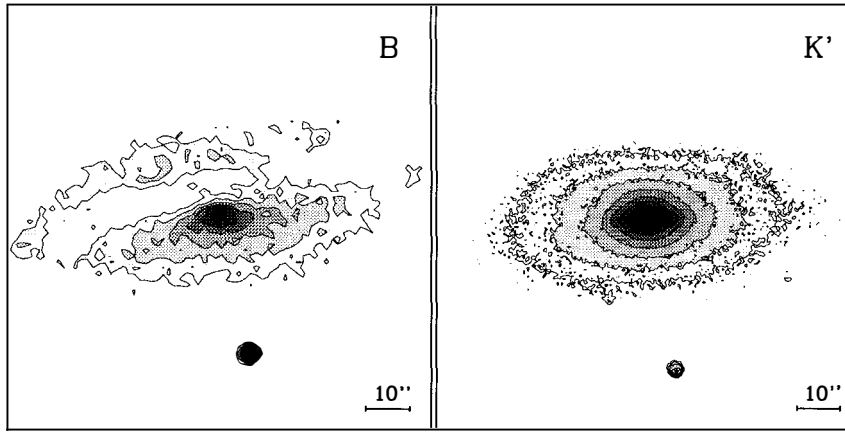


Figure 1. Galaxy NGC 7172, seen in blue (*left*) and NIR (*right*) light.[1)] Object at bottom is a foreground star.

2) The second difference is that, contrary to NIR light, optical light is extremely sensitive to populations of newborn stars, because the more massive (and hot) stars are extremely luminous, and are the only very blue stars (for cooler stars, the blue band is in the Wien part of the spectrum). As stars exhaust their nuclear hydrogen and evolve to the red giant branch, they become substantially redder. The more massive ones evolve to cool Red Giants in very short times (a few million years) and continue their evolution and finally blow up as supernovae in even shorter times. The NIR band picks up principally cool stars, while the hot, massive, and young ones, moderately visible as the NIR lies in the Rayleigh-Jeans part of their spectrum, are too rare to contribute significantly to NIR light. As a result, blue light tends to pick up galaxies with very recent star formation, while NIR light is less affected by very recent star formation and traces better the *stellar mass* content of galaxies.

However, NIR light *is* affected by star formation occurring roughly 10 million years

ago, because after 10 million years the typical very massive stars are in their Red Giant or subsequent phases before going supernova. Nevertheless, averaged over all epochs of star formation, the NIR wavelength domain seems to be the optimal one to avoid the effects of star formation.

These differences between optical light and NIR light have several important implications:

1) The lack of extinction allows one to probe the large-scale distribution of galaxies behind the dust-filled plane of our own Galaxy.[2)]

2) The galaxy tracers of the underlying matter distribution in the Universe is less biased by recent star formation in the NIR than in other bands.

2. The DENIS and 2MASS Surveys

With these advantages in mind, two consortia have set out to map the sky at NIR wavelengths: *DENIS* (DEep Near Infrared Southern Sky Survey) and *2MASS* (Two Micron All Sky Survey) The characteristics of the two surveys are given in Table 1.

Table 1: The DENIS and 2MASS surveys

	DENIS	2MASS
Institutions	Meudon, IAP, Leiden ...	UMass, IPAC, CfA ...
Hemispheres	South ($\delta < 2°$)	North + South
Telescopes	ESO 1m	2 new 1.3m
Color bands	I (0.8 μm), J (1.25 μm), K_s (2.15 μm)	J, H (1.65 μm), K_s
Detector	CCD 1024^2 (I), NICMOS-3 256^2 (JK)	NICMOS-3 256^2
Pixel size	$1''$ (I), $3''$ (JK)	$2''$
Quantum efficiency	0.4 (I), 0.65 (JK)	0.65
Exposure time	8 s (I), 9×1 s (JK)	6×1.3 s
Read-out noise	$8\,e^-$ (I), $30\,e^-$ (JK)	$30\,e^-$
Observing mode	Stop & Stare	Freeze-Frame Scan
Scan geometry	$12' \times 30°$	$8' \times 6°$
Survey years	1996–2000	1997–2000 (N), 1998–2000 (S)
Primary data	4000 GBytes	19 000 GBytes
Cost	\$3 million	\$30 million

Notes: The NIR exposures are dithered for better angular resolution.

3. Cosmological Applications

The applications of NIR surveys such as DENIS and 2MASS for extragalactic astronomy and cosmology have been described elsewhere[3,4,5,6,7)] and are briefly outlined again here.

3.1 Statistical properties of NIR galaxies

The large sample sizes ($> 10^4$, 10^5, and 10^6 galaxies in K, J, and I, respectively, for DENIS[5]) will help study correlations between properties.

3.2 Two-dimensional structure of the local Universe.

Catalogs of groups and clusters will be obtained from the galaxy lists extracted from the survey images. Statistical measures of large-scale structure will be obtained (such as the angular correlation function and higher order functions, counts in cells, and topological measures), in particular statistics on the full sphere from 2MASS data. Not only are NIR bands cleaner than optical bands to study the large-scale distribution of galaxies in the Universe (because less sensitive to recent star formation), but the multi-color aspect of the DENIS and 2MASS surveys allows one to see how all this 2D structure will vary with waveband, and indicate possible biases when going from one waveband to another. The alternative is that any such difference in 2D structure may be a reflection of selection effects, but we are working hard on avoiding this. Moreover, NIR surveys will probe the largest local concentration of matter, the *Great Attractor*, which is situated roughly right behind the Galactic Plane[8], and contains the closest rich cluster of galaxies.[9] Unfortunately, although extinction is beaten at low galactic latitudes, confusion with stars becomes a serious issue, for example on the accuracy of the photometry of fairly bright galaxies.[10]

3.3 Color segregation

Instead of studying structure versus waveband, one can study the inverse problem of understanding colors as a function of structure, hence environment. Color segregation is a potentially powerful probe of three dimensional morphological segregation of galaxies in the Universe (*e.g.,* the fact that the cores of galaxy clusters have the highest fraction of galaxies with elliptical morphological types).

3.4 Normalization of galaxy counts at the bright-end

Counting galaxies as a function of apparent magnitude provides better results at the faint-end than at the bright-end, simply because the bright-end suffers from very poor statistics (in the local uniform Universe, galaxy counts rise roughly as $\mathrm{dex}[0.6\,m]$). If our Local Group sits in an underdense region, we should see a lack of galaxies at the very bright end of the galaxy counts, which is brighter than the DENIS and 2MASS complete/reliable extraction limits. However, the error bars in recent studies[11,12,13] are too large to draw firm conclusions, and very wide-angle surveys such as DENIS and 2MASS will bring them down.

3.5 Mapping interstellar extinction

Among the many ways one can map the interstellar extinction in our Galaxy, one is to use galaxy counts[14,15], since the count normalization is shifted downwards when galaxies are extinguished. Galaxy colors may provide better extinction estimates.[6]

3.6 Cosmic Dipole

With its full sky coverage, 2MASS will be able to probe the *cosmic dipole*, which computes the vector sum of the flux *vectors* from the detected extragalactic objects, and which should be close to the peculiar acceleration of our galaxy, assuming that galaxy NIR light is a good tracer of the total mass content of the Universe. In particular, it will be interesting to compare the cosmic dipole with that obtained with the sparser IRAS galaxy samples.

4. A spectroscopic followup

There is much to be gained from knowing the third dimension in any galaxy survey. In particular, the statistics of the 3D galaxy distribution can be studied, and projection effects are virtuaally eliminated, although the 3D statistics are messed up by peculiar velocities. 3D color segregation can be studied as well as the convergence of the cosmic dipole with distance. And finally one gains access to the internal kinematics of structures.

For these reasons, a spectroscopic followup of the DENIS and 2MASS samples are highly desirable. In particular, DENIS will extract with high completeness and reliability over $\simeq 160\,000$ galaxies at $J < 14.4$.[5] Only 25 000 galaxies are expected in the largest complete and reliable K-band sample[4], although if cooling of the DENIS optics is implemented this Autumn, as scheduled, this number could be multiplied by three to four. The following discussion attempts to optimize the time for obtaining the largest complete J selected spectroscopic sample.

4.1 Simulated galaxy samples

To begin, galaxy samples are simulated with random positions in a uniform Universe, random blue galaxy luminosities from a Schechter[16] luminosity function, random bulge/disk ratio, disk inclination, and galactic latitude, plus scatter in the surface brightness versus luminosity relations for bulges and disks. Working separately on bulges and disks, the blue and NIR apparent magnitudes of each component are estimated using standard colors, k-corrections (the effect of redshifting a spectrum through a fixed observation wavelength filter), but no luminosity evolution. Samples of typically 25 000 galaxies are simulated with $z < 0.25$ and $L > 10\,L_*$, and subsamples of typically a few thousand galaxies are extracted with apparent magnitude limits in the blue or the NIR.

4.2 Mean galaxy surface brightness within fiber apertures

The surface brightness of galaxies are well represented by $\Sigma = \Sigma(0)\exp[(-r/r_1)]^\beta$, where $\beta = 1$ for exponential disks and $\beta = 1/4$ for $r^{1/4}$ bulges. The mean surface magnitude within a circular aperture centered on the galaxy is then

$$\langle\mu\rangle = \mu_0 - 2.5\log\{2[1-(x+1)\exp(-x)]/x^2\}\ , \tag{1}$$

for exponential disks and

$$\langle\mu\rangle = \mu_0 - 2.5\log(8/[b^8x^2]) - 2.5\log\gamma(8, bx^{1/4})\ , \tag{2}$$

for $r^{1/4}$ bulges. Here, $b = 7.67$ and

$$\gamma(8, bx^{1/4}) = 5040 - \exp(-bx^{1/4}) \sum_{k=0}^{7} \frac{7!}{k!} (bx^{1/4})^k \ , \tag{3}$$

and, in both cases, $x = \theta_{\rm fib}/(2\theta_1)$, with the angular radius $\theta_1 = \mathrm{dex}[-0.2(m - \mu_0)]/c^{1/2}$ is the angular scale length for exponential disks and the angular effective radius for $r^{1/4}$ bulges, and where $c = 2\pi$ for exponential disks and $c = 8!\,\pi/b^8 = 0.0106$ for $r^{1/4}$ bulges. We correct the mean surface brightness for inclination and internal extinction of the disk, for Galactic extinction, and for k-corrections and cosmological dimming ($\Sigma \sim (1+z)^{-4}$).

Because spectroscopy of nearby galaxies is done in the optical (we know very little of galaxy spectra in the NIR), one has to be careful when making optical observations of a NIR selected sample. Figure 2 shows the optical surface magnitudes (eqs. [1-3], corrected as explained above) versus NIR magnitude for a $J < 14.1$, subsample and illustrates that the lowest surface brightness galaxies (which are the most difficult to obtain spectra for, see below) are distant ellipticals.

4.3 Time constraints for different instruments

The number of spectroscopic nights required to achieve a complete J-limited sample in the southern hemisphere has been estimated for different telescopes in the Southern hemisphere with additional Monte-Carlo simulations. The mean surface magnitudes of galaxies, computed over different fiber sizes, are converted to observing times for continuum $S/N = 5$, *assuming that all telescopes have the same transmission from sky to detector*, and normalizing to the published[18] performance of FLAIR-II, a 92-fiber spectroscope on the very wide-field (6° Schmidt) UKST telescope in Australia. The maximum time to reach $S/N = 5$ is recorded for each set of N_f galaxies (where N_f is the number of fibers), after the galaxies have been first sorted by galactic latitude.

Figure 3 shows the comparison of the different available (or soon to be) or potential (FIFI, FLAIR IIIa,b,c) instruments. Only 60 to 85 nights are required for an 80 000 galaxy sample limited to $J < 13.7$ on an upgrade of FLAIR II, with more fibers and automatic configuration. This would correspond to one year at 50% pressure on the UKST telescope (which is also used for Schmidt plate photographic surveys).

Discussions are currently underway for upgrading FLAIR II and performing a spectroscopic followup to DENIS. These involve Q. Parker and M. Colless in Australia, W. Saunders in Edinburgh, and V. Cayatte, H. Di Nella, R. Kraan-Korteweg, G. Paturel and others in France. Our target is to complete the spectroscopic survey in 2000 or 2001, starting towards the end of the DENIS survey.

References

1] Héraudeau, P., Simien, F. & Mamon, G.A., 1996, *A&AS*, **117**, 417.

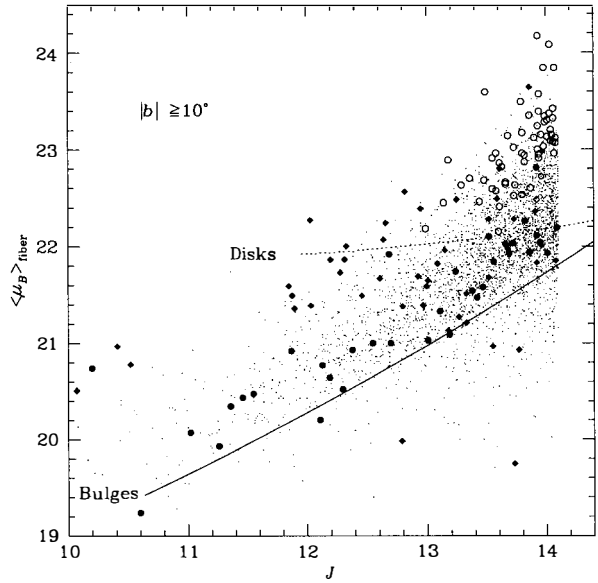


Figure 2. Simulation of surface magnitudes of galaxies, averaged over FLAIR 6.7″ fibers, as a function of NIR J magnitude for an NIR selected selected sample of 3300 galaxies, with galactic latitude $|b| > 10°$. Late-type spiral galaxies ($\geq$ 90% of disc) are shown as *diamonds* and elliptical galaxies ($\geq$ 90% of bulge) by *circles*. The *full symbols* correspond to nearby galaxies ($z \leq 0.03$) and the *open symbols* to distant galaxies ($z \geq 0.1$). Galaxies with intermediate bulge/disc ratio, as well as galaxies at intermediate redshift are shown as *points*. The *curves* represent the theoretical relations for nearby galaxies (for bulges we have adopted a luminosity dependent central surface magnitude[17]), with colors corresponding to a reddening at the galactic pole of 0.17. Discs and bulges above their respective theoretical curves have greater reddening, cosmological dimming, and k-correction. Discs below the theoretical disc curve are inclined.

2] Mamon, G.A., 1994, *in "Unveiling Large-Scale Structures behind the Galactic Plane", ASP vol. 67*, ed. C. Balkowski & R.C. Kraan-Korteweg (San Francisco: ASP), p. 53.

3] Mamon, G.A., 1994, *ApSS*, **217**, 237.

4] Mamon, G.A., 1995, *in 35th Herstmonceaux Conference: "Wide-Field Spectroscopy and the Distant Universe"*, ed. S.J. Maddox & A. Aragón-Salamanca (Singapore: World Scientific), p. 73.

5] Mamon, G.A., 1996, *in "Spiral Galaxies in the Near-IR"*, ed. D. Minniti & H.-W. Rix (Berlin: Springer), p. 195.

6] Mamon, G.A., Banchet, V., Tricottet, M. & Katz, D., 1997, *in "The Impact of Large Scale Near Infrared Surveys"*, ed. F. Garzón (Dordrecht: Kluwer), in press.

7] Schneider, S., 1997, *in "The Impact of Large Scale Near Infrared Surveys"*, ed. F. Garzón (Dordrecht: Kluwer), in press.

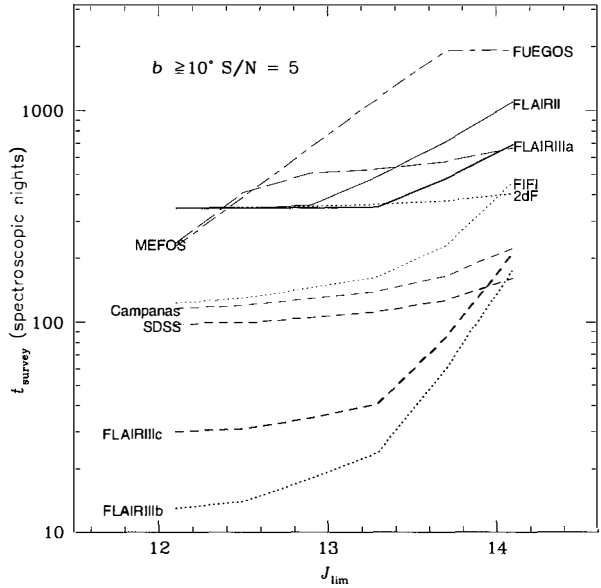


Figure 3. Estimates of time in spectroscopic nights to complete *without loss of galaxies* a survey with $|b| \geq 10°$ and $S/N \geq 5$, using mean surface brightnesses within fibers, for the simulated galaxy sample. FLAIR IIIa would have 150 fibers, with slow day-time configuration. FLAIR IIIb and IIIc would have 250 fibers with respectively 30 and 10 minute on-telescope configuration. FIFI is a copy of MEFOS that could easily be mounted at the SAAO in South Africa.

8] Kolatt, T., Dekel, A. & Lahav, 1995, *MNRAS*, **275**, 797.
9] Kraan-Korteweg, R.C., Woudt, P.A., Cayatte, V., Fairall, A.P., Balkowski, C. & Henning, P.A., 1996, *Nature*, **379**, 519.
10] Chester, T., 1997, *in "The Impact of Large Scale Near Infrared Surveys"*, ed. F. Garzón (Dordrecht: Kluwer), in press.
11] Mobasher, B., Sharples, R.M. & Ellis, R.S., 1986, *MNRAS*, **223**, 11.
12] Gardner, J.P., Sharples, R.M., Carrasco, B.E. & Frenk, C.S., 1996, *MNRAS*, in press.
13] Huang, J.-S., Cowie, L.L., Ellis, R.S., Gardner, J.P., Glazebrook, K., Hu, E.M., Songaila, A. & Wainscoat, R.J., 1996, submitted to *ApJ*.
14] Burstein, D. & Heiles, C., 1978, *Ap. Lett.*, **19**, 69.
15] Burstein, D. & Heiles, C., 1982, *AJ*, **87**, 1165.
16] Schechter, P.L., 1976, *ApJ*, **203**, 297.
17] Sandage, A. & Perelmuter, J.-M., 1990, *ApJ*, **361**, 1.
18] Parker, Q.A., 1995, *Spectrum*, **7**, 17.

THE COBRAS/SAMBA SPACE PROJECT

N.MANDOLESI[1], M.BERSANELLI[2], L.DANESE[3], DE ZOTTI[4] AND N.VITTORIO[5]

[1]*Istituto TESRE, CNR, Bologna, Italy*
[2]*Istituto di Fisica Cosmica, CNR, Milano, Italy*
[3]*International School for Advanced Studies, Trieste, Italy*
[4]*Osservatorio Astronomico di Padova, Padova Italy*
[5]*Dipartimento di Fisica, Universita' di Tor Vergata, Roma, Italy*

ABSTRACT — The ESA mission COBRAS/SAMBA is designed to produce extended, accurate, multifrequency maps of the anisotropy of the Cosmic Microwave Background (CMB) radiation, with angular sensitivity from 0.1 to 180 degrees. This will allow a full identification of the primordial density perturbations which grew to form the large–scale structures observed in the present universe. The COBRAS/SAMBA observations will provide decisive tests of competing theories describing the origin of large-scale structure, and will lead to very accurate determination of all cosmological parameters. A combination of bolometric and radiometric detection techniques will ensure the sensitivity and wide spectral coverage required for accurate foreground discrimination. The project has recently completed the feasibility study within the European Space Agency M3 programme[1].

1 Introduction

Remarkable progress has been achieved in cosmology in recent years, regarding both new observational results and consolidation of the underlying theory. New, accurate observations of the Cosmic Microwave Background (CMB) have been a major component of such progress. In 1992 the COBE team announced the detection of intrinsic temperature fluctuations in the CMB at angular scales $\theta > 7°$, with brightness amplitude $\Delta T/T \sim 10^{-5}$ (Smoot et al. 1992). Shortly after the COBE discovery, experiments from sub–orbital platforms have confirmed the detection of large scale CMB anisotropies at comparable $\Delta T/T$ amplitudes (Ganga et al 1993, Hancock et al 1994). Fluctuations at these angular scales have been interpreted as due to the Sachs–Wolfe effect, i.e., gravitational red–shifting and blue–shifting of the CMB photons from the last scattering surface (Sachs & Wolfe 1967). The COBE results are consistent with a scale–invariant primordial spectrum, as predicted by most inflationary scenarios, placing constraint on the spectral index n at about 30% level.

The angular resolution of the COBE Differential Microwave Radiometer (DMR) provides information on the CMB fluctuation power spectrum only up to multipoles

[1]Before the publication of this work COBRAS/SAMBA has been recommended by ESA as the next scientific medium–size mission M3 (May 1996).

$\ell \sim 20$, leaving completely untouched the wealth of cosmological information encoded in the shape of the fluctuation power spectrum at higher multipoles. As it will be outlined in the next section, an accurate reconstruction of the CMB fluctuation power spectrum would represent a major breakthrough for cosmology and astro–particle physics. Several groups have successfully carried out ground based and balloon–borne anisotropy experiments at sub–degree angular scales to address the problem (see Scott et al. 1995 for a review). However, it is clear that only a dedicated space mission can ensure the mapping capability required to reconstruct the entire power spectrum with high precision (Danese et al. 1996).

The recognition of this unique scientific opportunity has lead a number of European groups to undertake a joint effort within the ESA M3 program to plan a second generation space experiment dedicated to measurements of the CMB anisotropies. The COBRAS/SAMBA mission is designed to produce *definitive* maps of the CMB, in the sense that they will be limited only by unavoidable astrophysical limits rather than instrumental noise or systematic effects. This ambitious goal requires to image nearly the whole sky with an angular resolution $\sim 10'$ and with a sensitivity approaching $\Delta T/T \sim 10^{-6}$ per resolution element. In addition, a wide spectral range (30–1000 GHz) is necessary to safely disentangle foreground emissions from the anisotropies of cosmic origin. COBRAS/SAMBA is the result of the merging of two proposals presented in 1993 to the European Space Agency *M3 Call for Mission Ideas*: COBRAS (Cosmic Background Radiation Anisotropy Satellite; Mandolesi et al. 1993) and SAMBA (Satellite for Measurements of Background Anisotropies; Puget et al. 1993). The COBRAS/SAMBA team completed the ESA feasibility study in early 1996, and the project will be presented in April 1996 for the final ESA selection[1].

2 Scientific Objectives

COBRAS/SAMBA will produce near all-sky maps in 9 frequency bands in the range 30–1000 GHz, providing a high resolution imaging of the last scattering surface. The CMB temperature fluctuations are often represented in a spherical harmonic decomposition:

$$\frac{\Delta T}{T}(\theta,\phi) = \sum_{\ell,m} a_\ell^m Y_\ell^m(\theta,\phi). \qquad (1)$$

The fluctuations power spectrum is defined by the mean square of the coefficients a_ℓ^m.

$$C_\ell = \frac{1}{2\ell+1} \sum_{m=-\ell}^{\ell} |a_\ell^m|^2. \qquad (2)$$

COBRAS/SAMBA will allow a measure of all multipoles C_ℓ of the temperature anisotropies from $\ell = 1$ (dipole term) up to $\ell \simeq 1200$ (corresponding to a beam with FWHM$\simeq 7'$). Different cosmological scenarios produce different power spectra. The full, accurate reconstruction of the C_ℓ spectrum is the main observational goal of the mission, and it will set the basis for a new, fundamental study of the early universe.

The COBRAS/SAMBA observations at high angular resolution will provide key tests for structure formation mechanisms, based on the statistics of the observed $\Delta T/T$

distribution. Inflationary models predict Gaussian fluctuations, while alternative models based on the presence of topological defects, such as strings, monopoles, and textures, predict non–Gaussian statistics (e.g. Turok 1989, Coulson et al. 1994). The angular resolution and sensitivity of COBRAS/SAMBA will allow discrimination between these alternatives.

Accurate determination of key cosmological parameters (such as Ω, Ω_b, H_0, Λ) can be achieved if the power spectrum is measured with very high precision up to $\ell \sim 1200$ (i.e., both the main Doppler peak, at $\ell \simeq 220\sqrt{\Omega_0}$, and the secondary peaks at higher ℓ). This will be the case for COBRAS/SAMBA, thanks to its high angular resolution, and the wide frequency range to control foregrounds contamination. Simulations[2] have shown that these parameters can be determined *independently of each other* with an accuracy of a few percent. For example, it can be shown that the observations of COBRAS/SAMBA will lead to a determination of the total density parameter Ω_0 and of the Hubble constant H_0 with 1% accuracy, while the baryon density Ω_b can be measured with $\sim 3\%$ accuracy. These precisions are far beyond what can be expected from conventional astrophysical methods in the foreseeable future. These accuracies can be achieved only with very careful handling of the foregrounds, and they are strong functions of the angular resolution. For an experiment with $30'$ angular resolution and with the same sensitivity per pixel of COBRAS/SAMBA, the accuracy on these parameters would be degraded by a factor of 5 to 8.

The COBRAS/SAMBA maps will provide fundamental informations also at large angular scales, where the expected sensitivity is much greater than the one of COBE–DMR. The large scale fluctuations will allow an accurate measure of the spectral index n of the primordial fluctuation spectrum:

$$(\delta\phi)^2 \propto \lambda^{(1-n)}, \tag{3}$$

where $\delta\phi$ is the potential fluctuation responsible for the CMB anisotropies, and λ is the scale of the density perturbation. This corresponds to CMB temperature fluctuations $(\Delta T/T)^2 \propto \theta^{(1-n)}$ for angles greater than the horizon scale θ_h, i.e., $\theta \gtrsim \theta_h \simeq 2^\circ\ \Omega_0^{1/2}$. The proposed observations will be able to verify the scale invariant spectrum ($n = 1$) predicted by most inflation theories with an accuracy $\sim 3\%$. For comparison, the COBE–DMR limit on the spectral index are of order 30–40% (Bennett et al. 1994, Gorski et al. 1994).

Temperature anisotropies at large angular scales can be generated by gravitational waves (tensor modes), in addition to the energy-density perturbation component (scalar modes) whose fluctuations are characterized by spectral indexes n_t and n_s, respectively. The parameters n_t and n_s and the ratio $r \equiv C_2^t/C_2^s$ between the CMB quadrupoles ($\ell = 2$) from gravitational waves and density perturbation components, can be measured by COBRAS/SAMBA, thus providing information on the effective potential which controls inflation. Testing inflation with observations of the CMB anisotropies is one

[2]See the COBRAS/SAMBA Report on the Phase A study, M.Bersanelli, F.Bouchet, G.Efstathiou, M.griffin, J.M.Lamarre, N.Mandolesi, H.U. Norgaard–Nielsen, O.Pace, J.Polny, J.L.Puget, J.Tauber, N.Vittorio, S.Volonte; ESA D/SCI(96)3 February 1996. Requests for copies of this report should be addressed to J.Tauber, at ESTEC, Keplerlaan 1, P.O.Box 299, 2200 AG Noordwijk, The Netherlands.

of the few ways of testing physics at ultra–high energies levels ($\sim 10^{15}$ GeV) which can't be reached by any foreseeable particle–accelerators.

COBRAS/SAMBA will also measure the Sunyaev–Zel'dovich effect for more than 1000 rich clusters, using the high resolution bolometric channels. Combined with X–ray observations these measurements can be used to yield an independent estimate the Hubble constant H_0 (Cavaliere, Danese & De Zotti 1977, 1979, Silk & White 1978). The product maps will also constitute a major data base relevant to many aspects of astronomy and astrophysics. The multi-frequency observations will produce large catalogs of IR and radio galaxies, AGNs, QSOs, Blazars. The COBRAS/SAMBA data base will be used also for studies of our Galaxy through the mapping of its diffuse emission. Dust properties can be investigated, and the morphology of interstellar clouds and cirrus can be mapped. The synchrotron and free-free emissions at low frequencies will be studied, leading to new determination of the (spatially dependent) spectral indexes, with important impact on the study of cosmic ray and magnetic field distributions.

3 Foreground Radiation

The COBRAS/SAMBA mission is designed to recover the power spectrum with an accuracy which is limited by cosmic variance and unavoidable astrophysical foregrounds. In fact, in addition to the CMB temperature fluctuations, foreground structures will be caused by weak, unresolved extragalactic sources and by radiation of galactic origin (interstellar dust, free–free and synchrotron radiation). The COBRAS/SAMBA observations will reach the required control on the foreground components in two ways. First, the large sky coverage (> 90%) will allow accurate modeling of these components where they are dominant (e.g. Galactic radiation near the galactic plane). Second, the observations will be performed in a spectral range as broad as possible. The COBRAS/SAMBA channels will span the spectral region of minimum foreground intensity (60–200 GHz), but with enough margin at high and low frequency to monitor the effect of the various foreground components. By using the COBRAS/SAMBA spectral information it will be possible to remove the foreground contributions with high accuracy[2].

The expected fluctuation level at angular resolutions of 10′ and 30′ due to Galactic emission and extragalactic sources is shown in Figure 1. In the high frequency channels ($\nu \gtrsim 140$ GHz) the main foreground components will be dust emission from infrared cirrus and from normal spiral galaxies, as well as that from starburst galaxies; towards the low–frequency range ($\nu \lesssim 70$ GHz) galactic synchrotron and free–free emission become increasingly important (Toffolatti et al. 1995). Calculations of the residual temperature fluctuations from unresolved extragalactic sources show that in the range 50–300 GHz they will contribute $\Delta T/T \lesssim 10^{-6}$, i.e. below the expected noise level. The final limitation to the cosmological information of the COBRAS/SAMBA maps is expected to be due to the residual uncertainties in the separation of the foreground components rather than statistical noise. Therefore, the overall design of the instrument and payload is highly driven by the need of achieving a spectral coverage

as large as possible. Performing measurements where the dominant foreground components are different will permit a powerful cross check on residual systematic errors in the CMB temperature fluctuation maps.

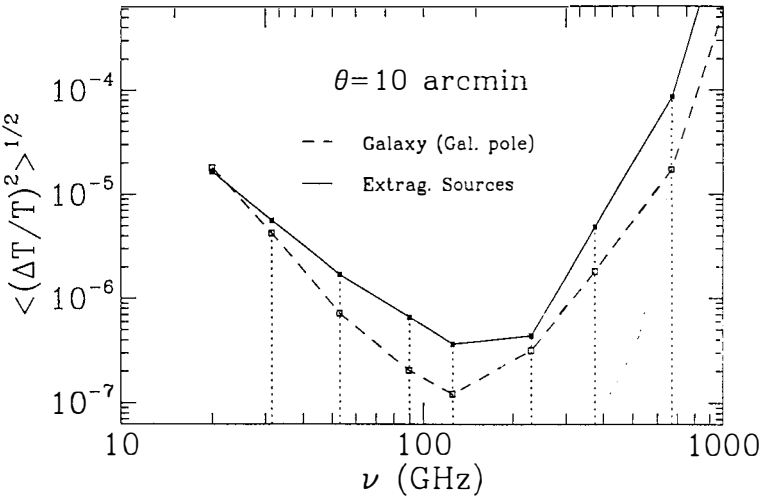


Figure 1 – Estimated fluctuation levels due to Galactic polar emission (dashed lines) and to extragalactic sources (solid lines) assuming 10′ angular resolution (see Toffolatti et al. 1995).

4 The Model Payload

The payload consists mainly of a shielded, off-axis telescope, leading to an integrated instrument (Focal Plane Assembly, FPA). The payload is part of a spinning spacecraft, with a spin rate of 1 rpm. Figure 2 shows a sketch of the payload, with the main features of the optical system and of the FPA. The focal plane assembly is divided into low-frequency (LFI) and high-frequency (HFI) instrumentation according to the technology of the detectors. Both the LFI and the HFI are designed to produce high-sensitivity, multifrequency measurements of the diffuse sky radiation. The LFI will measure in four bands in the frequency range 30–130 GHz (2.3–10 mm wavelength). The HFI will measure in five channels in the range 140–1000 GHz (0.3–2.1 mm wavelength). The highest frequency LFI channel and the lowest HFI channel overlap near the minimum foreground region. Table 1 summarizes the main characteristics of the COBRAS/SAMBA payload.

Table I. Payload Characteristics

Telescope	1.5 m Diam. Gregorian; system emissivity 1% Viewing direction offset 70° from spin axis								
Instrument	LFI				HFI				
Detector Technology	HEMT receiver arrays				Bolometer arrays				
Detector Temperature	~ 100 K				0.1 - 0.15 K				
Cooling Requirements	Passive				Cryocooler + Dilution system				
Center Frequency (GHz)	31.5	53	90	125	143	217	353	545	857
Wavelength (mm)	9.5	5.7	3.3	2.4	2.1	1.38	0.84	0.55	0.35
Number of Detectors	4	14	26	12	8	12	12	12	12
Angular Resolution (arcmin)	30	18	12	12	10.3	7.1	4.4	4.4	4.4
Optical Transmission	1	1	1	1	0.3	0.3	0.3	0.3	0.3
Bandwidth ($\frac{\Delta\nu}{\nu}$)	0.15	0.15	0.15	0.15	0.37	0.37	0.37	0.37	0.37
$\frac{\Delta T}{T}$ Sensitivity per pixel (14 months, 1σ, 10^{-6} units)	7.8	7.5	14.4	35.4	1.2	2.0	12.1	76.6	4166

An off-axis, Gregorian telescope will be used, to avoid spurious signals arising from the mirrors or from supports and mechanical mounting. The primary parabolic mirror has a diameter of approximately 1.5 meter, with an elliptic subreflector of 0.57 m. Stray satellite radiation and other off–axis emissions are minimized by underilluminating the low–emissivity optics. The telescope reimages the sky onto the focal plane instrument located near the payload platform. Blockage is a particularly important factor since several feeds and detectors are located in the focal plane, and unwanted, local radiation (e.g. from the Earth, the Sun and the Moon) needs to be efficiently rejected. A large, flared shield surrounds the entire telescope and FPA, to screen the detectors from contaminating sources of radiation. The shield also plays an important role as an element of the passive thermal control of the spacecraft.

The necessary wide spectral range requires the use of two different technologies, bolometers and radiometers, incorporated in a single instrument. Both technologies have shown impressive progress in the last ten years or so, and more is expected in the near future. The thermal requirements of the two types of detectors are widely different. The coherent radiometers (LFI), operating in the low frequency channels, give good performance at operational temperature of ~ 100 K, which is achievable with passive cooling. The bolometers, on the other hand, require temperatures of 0.1 to 0.15 K in order to reach their goal performances. The LFI consists of an array of 28 corrugated, conical horns, each exploited in the two orthogonal polarization modes, feeding a set of state–of–the–art, high sensitivity radio receivers. The receivers will be based on MMIC (Monolithic Microwave Integrated Circuits) technology with HEMT (High Electron Mobility Transistor) ultra–low noise amplifiers (see e.g. Pospieszalski 1993). Since the whole LFI system will be passively cooled, it can be operated for a duration limited only by spacecraft consumables (up to 5 years). Fifty-six bolometers will be used for the HFI array, which require cooling at ~ 0.1 K. The cooling system combines active coolers reaching 4 K with a dilution refrigeration system working in

zero gravity. The refrigeration system will include two pressurized tanks of ^{3}He and ^{4}He for an operational lifetime of 2 years.

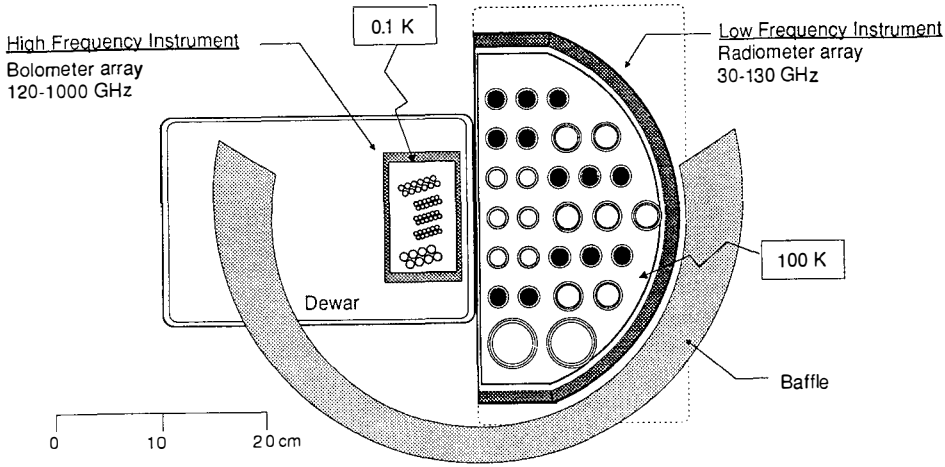


Figure 2 – A schematic top view of the Focal Plane Assembly, with the bolometric system (HFI) cooled at 0.1 K, sharing the focal plane with the arrays of passively cooled feed horns (LFI).

5 Orbit and Sky Observation Strategy

The COBRAS/SAMBA mission plan is designed for maximum control over possible residual systematic effects. One of the main requirements is to observe from a far-Earth orbit with a suitable scan geometry in order to avoid large solid-angle planetary sources. In fact, even from an orbit at lunar distance stray radiation from the Earth, the Sun and the Moon is a serious potential contaminant (Muciaccia et al. 1996). COBRAS/SAMBA will operate from a Lissajous orbit around the L2 Lagrangian point of the Sun–Earth system, at a distance of about 1,500,000 km from the Earth. From this nearly-ideal location the required off-axis rejection is relaxed to a level which is achievable with careful, standard optical designs. The selected L2 orbit is also very favorable from the point of view of passive cooling and thermal stability. The spacecraft will be normally operated in the anti-solar direction, with part of the sky observations performed within $\pm 15°$ from anti-sun.

The telescope optical axis is offset by 70° from the spin axis. Thus at each spacecraft spin rotation the telescope pointing direction sweeps a 140° circle in the sky. This scan strategy provides proper redundancy for checking the effect of possible instrumental drifts, and allows an efficient absolute calibration exploiting the CMB dipole (as measured by COBE-DMR) and the spacecraft orbital velocity (Bersanelli et al. 1996).

The accuracy of the overall calibration is expected to be $< 0.2\%$.

6 Conclusions

An accurate, extensive imaging of the CMB anisotropies with sub–degree angular resolution would provide decisive answers to several major open questions on structure formation and cosmological scenarios. The observational requirements of such an ambitious objective can be met by a space mission with a far–Earth orbit and instruments based on state–of–the–art technologies. While COBRAS/SAMBA is a mission dedicated to cosmological research, it will provide an accurate, multifrequency survey in a spectral range extremely interesting for a wide variety of astrophysical problems.

The Phase A study has shown that a mission combining bolometric and radiometric technologies is feasible within the European Space Agency medium-size mission scope. The wide spectral range (30–800 GHz) and the sensitivity to a very broad spectrum of angular scales ($\sim$ 10′ to 180 degrees), are distinctive features of COBRAS/SAMBA. Simultaneous observations of the CMB anisotropies in bands characterized by different foreground emissions and independent instrumental techniques make the COBRAS/SAMBA mission extremely robust against any possible residual systematic effects.

References

Bennett, C.L., et al. 1994, ApJ, 436, 423.

Bersanelli, M., Natoli, P., Muciaccia, F., Vittorio, N., and Mandolesi, N., 1996, A&A Suppl., in press.

Cavaliere, A., Danese, L., & De Zotti, G., 1977, ApJ 217, 6.

Cavaliere, A., Danese, L., & De Zotti, G., 1979, A&A 75, 322.

Coulson, D., Ferreira, P., Graham, P., & Turok, N., 1994, Nature, 368, 27

Danese, L., Toffolatti, L., Franceschini, A., Bersanelli, M., Mandolesi, N., 1995, Astroph. Lett. & Comm., 33, 257.

Ganga, K., et al. 1993, ApJ, 410, L57.

Gorski, K., Hinshaw, G., Banday, A.J., Bennett, C.L., Wright, E.L., Kogut, A., Smoot, G.F., & Lubin, P.M. 1994, ApJ, 430, L89.

Hancock, S., et al. 1994, Nature, 367, 333.

Mandolesi, N., et al. 1993, *COBRAS* proposal for ESA M3.

Muciaccia, F., et al. 1996, A&A Suppl., in press.

Pospieszalski, M.W. 1993, 23rd EuMC, Sept 1993, p.73.

Puget, J.L., et al. 1993, *SAMBA* proposal for ESA M3.

Sachs, R.K. & Wolfe, A.M. 1967, ApJ 147, 73

Scott, D., Silk, J., & White, M., 1995, Science, 268, 829.

Silk, J. & White, S. 1979, ApJ 226, L103.

Smoot, G.F., et al, 1992, ApJ, 396, L1.

Toffolatti, L., et al., 1995, Astrophys. Lett. & Comm., 32, 125.

Turok, N. 1989, Phys. Rev. Lett. 63, 2625.

WHAT CAN WE STILL LEARN FROM THE COSMIC BACKGROUND RADIATION ?

Jacques Delabrouille
Institut d'Astrophysique Spatiale, CNRS & Université Paris XI, bât 121, 91405 Orsay Cedex, France
and
Enrico Fermi Institute, University of Chicago, 5460 South Ellis Avenue, Chicago, IL 60510, USA

1 Introduction

After the encouraging results of the COBE satellite, there has been a burst of new interest in the anisotropies of the Cosmic Microwave Background Radiation (CMBR), both on the experimental side and on the theoretical side. While experimenters, in answer to announcements of opportunities by several space agencies, proposed sophisticated new generation satellites to map the fluctuations of the CMBR with a sensitivity and angular resolution an order of magnitude better than those of the COBE-DMR experiment, theorists refined their calculations to evaluate how individual parameters of the theories affect the expected properties of the tiny fluctuations of the CMBR. This paper is a short review of the main theoretical and experimental issues of this field, with a strong emphasis put on the scientific potential of the COBRAS/SAMBA satellite, proposed by a european collaboration in answer to the announcement of opportunity for the 3rd medium-sized ESA mission. The interested reader can get more details in the proceedings of the XXXIst Moriond meeting on Microwave Background Anisotropies (March 16-23, 1996), and in the COBRAS/SAMBA ESA report[1].

Before COBE revolutionized the field, the main concerns about the CMBR were whether its spectrum was indeed that of a blackbody, as predicted by the theory of the Big-Bang, and whether or not there were temperature fluctuations at a level sufficient to testify on the existence, at the time of last scattering, of the density fluctuations needed to seed structure formation. The first of these questions was answered by the FIRAS experiment (Mather *et al.*, 1993; Fixsen *et al.*, 1996) which, combined to a series of ground measurements, confirmed that the CMBR spectrum is indeed that of a blackbody, at a temperature of 2.738 ± 0.010 K, over frequencies ranging from 600 MHz to 600 GHz (see Smoot *et al.*, 1996, and references therein, for a review). The second question was answered to some extent by the successful detection of temperature anisotropies at the level of $\Delta T/T \sim 10^{-5}$ on large angular scales ($\sim 10°$) by the DMR experiment (Smoot *et al.*, 1992). Several ground-based and balloon-borne experiments have confirmed since then the existence of such levels of anisotropies in the temperature of the CMBR at large and intermediate angular scales (see e.g., White, Scott & Silk, 1994 for a review).

So, what can we still learn from the cosmic microwave background radiation, and how? In this paper, I first review briefly how the fluctuations of the CMBR are generated, and in which manner they are sensitive to the various cosmological parameters, assumptions and theories. Next, I address the important instrumental and observational issues that make the observation of CMBR anisotropies so difficult. Finally, I summarize the main characteristics and expected scientific outcome of the COBRAS/SAMBA mission.

[1] this report summarizes the results of the scientific and technical study activities of the COBRAS/SAMBA phase A; copies can be obtained from J. Tauber or O. Pace, both at ESTEC, Keplerlaan 1, P.O. Box 299, 2200 AG Noordwijk, The Netherlands, or from S. Volonté, Programme Coordination and Planning Office, ESA HQ, D/SCI, 8-10, rue Mario Nikis, 75738 Paris Cédex 15, France

2 Anisotropies of the CMBR

In addition to the dipole anisotropy of $3.343 \pm 0.016 mK$ (Smoot *et al.*, 1991), which can be interpreted as a consequence of our peculiar motion with respect to the radiation, and to higher order kinematic anisotropies, of much lower amplitude, anisotropies of the CMBR can be divided into two main categories. Primary anisotropies are generated on the Last Scattering Surface (LSS) at $z \sim 1100$ and before, while so-called secondary anisotropies are due to the interaction of CMBR photons with the matter constituting the Universe between $z = 1100$ and $z = 0$. The redshift of the LSS does not depend much on cosmological parameters, since the last scattering of photons depends essentially on the thermal history of the Universe (the Universe became transparent when most of the hydrogen recombined), which for the standard Big-Bang model with a Friedmann-Robertson-Walker metric is simply $T = T_0(1 + z)$.

2.1 Primary anisotropies

Primary anisotropies are initiated by some kind of initial perturbations to the homogeneous universe at very early times. Currently, the plausible scenarios are that either these primary perturbations are the result of quantum fluctuations in an inflationary epoch, or they are due to topological defects (monopoles, cosmic strings, domain walls or textures) generated during a phase transition in the early Universe. Both processes involve very high energy physics, well above the reach of existing particle accelerators. To these initial perturbations of the gravity field and of matter/particle/energy distribution correspond intrinsic temperature inhomogeneities, which depend on the nature and amplitude of the initial perturbations (see, e.g., Kolb & Turner, 1990).

These initial perturbations will then evolve due to gravity and particle interactions. In particular, before recombination, oscillations of the hot and dense plasma are expected to occur, depending on the models. These oscillations arise due to the driving force of gravity, resisted by the radiation pressure of relativistic particles tightly coupled to matter (as long as their mean free path is smaller than the scale of interest). Such oscillations, if present, will generate temperature anisotropies at the time of last scattering by the Sachs-Wolfe (gravitational redshift or blueshift of the photons coming out of an overdense or underdense region) and Doppler (due to matter motions) effects, and should be observed in the power spectrum[2] of temperature fluctuations in the form of one or more so-called "Doppler peaks" (which correspond to resonances for some values of ℓ, i.e., for some scales).

Given initial perturbations, it is possible to compute the evolution of these perturbations and the properties of the observed CMBR anisotropies. In most inflationary models, it is expected that the initial perturbations should be adiabatic fluctuations (pure density fluctuations, with no modification of the composition), randomly distributed according to a gaussian statistic, and with an initial power spectrum $P(k) \propto k^n$ ($n \simeq 1$). In this case, given a set of cosmological parameters (i.e. the values of H_0, Ω_0, Ω_b, Λ, ...)[3], it is possible to calculate the expected spectrum of the fluctuations with much accuracy (the coupled differential equations must be solved numerically, but the underlying physics is relatively simple and well understood, involving only linearized gravity and fundamental particle interactions). For models based on topological defects, the calculations are less simple, and the exact shape of the spectrum is not very well known yet. However, much work is in progress in this direction, and it is likely that similar predictions should be available before accurate and reliable data are available to distinguish between the models (see for instance Albrecht *et al.*, 1995, Crittenden & Turok, 1995, Coulson *et al.*, 1994).

[2] the power spectrum of the temperature fluctuations is defined by the values $C_\ell = \langle |a_{\ell m}|^2 \rangle_m$ for all integer ℓ, where the $a_{\ell m}$ are the coefficients of the spherical harmonic decomposition of the temperature fluctuations

$$\frac{\Delta T(\theta, \phi)}{T} = \sum_{\ell, m} a_{\ell m} Y_{\ell m}(\theta, \phi) \qquad (1)$$

in this decomposition large ℓ's correspond to small angular scales and small ℓ's to large angular scales

[3] H_0 is the Hubble constant, Ω_0 the ratio of total density to critical density, Ω_b the ratio of baryonic density to critical density, Λ the cosmological constant

If the inflationary paradigm favoured by most cosmologists is correct, the relative positions and heights of the peaks in the spectrum of the fluctuations bear unique signatures of the cosmological parameters (Hu & White, 1996; Kamionkowski *et al.*, 1994). For example, the position of the first Doppler peak is set by the maximum angular scale of the acoustic plasma oscillations at the time of last scattering and thus, since such oscillations cannot occur in causally unconnected regions, by the (sonic) horizon size at the LSS. But to first order, the angle subtended by the horizon at a redshift z depends only on the redshift and on the geometry of the Universe, that is, on Ω_0. Since the redshift of the LSS is almost independent of cosmological parameters, the position of the first Doppler peak gives a direct estimation of Ω_0 (through the dependance relation $\ell \sim 220 \times {\Omega_0}^{-1/2}$). Similarly, the relative heights of the Doppler peaks depend on Ω_b and H_0. In practice, it has been shown by a number of independent calculations (Efstathiou *et al.*, 1996, Jungman *et al.*, 1996) that in the standard inflationary models, most cosmological parameters can be recovered with percent level accuracy from a precise determination of the spectrum of the fluctuations. In other models, it is still expected that a lot of information is contained in the spectrum, but the exact accuracy that could be reached on cosmological parameters in defect scenarios, for instance, has not been determined yet.

2.2 Secondary anisotropies

Secondary anisotropies are generated after the last-scattering epoch. The most important source of secondary anisotropies is the Sunyaev-Zel'dovich effect, inverse Compton scattering of CMBR photons by free electrons as those present in the cores of clusters of galaxies. The SZ thermal effect, due to the interaction of CMBR photons with hot electrons with no bulk motion, generates a very characteristic distortion of the thermal CMBR spectrum, depleting the photon population in low energy states, augmenting it at higher energies (Zel'dovich & Sunyaev, 1969).

The kinetic SZ effect is due to the interaction of the CMBR photons with electrons having a bulk velocity with respect to the cosmological blackbody, and generates temperature fluctuations that are directly proportional to the radial velocity of the electrons, without any spectral deviation from a blackbody. Thus, fluctuations due to the kinetic SZ effect are undistinguishable from primary temperature fluctuations by their frequency spectrum alone. Fortunately, most ionized regions of the Universe that are susceptible to generate kinetic SZ effect by their peculiar motion (i.e. clusters of galaxies) can hopefully be identified from existing source catalogs and, if the electron gas is hot enough, from their thermal SZ signature. However, if there are colder ionized regions as for example in the case of an early reionization by quasars (Aghanim *et al.*, 1995), the thermal SZ effect might be negligible, and the kinetic SZ significant.

There is a number of other secondary effects, of much smaller amplitude, as the Rees-Sciama effect (Rees & Sciama, 1968) on photons crossing a time varying gravitational potential, the gravitational lensing effect on the CMBR, or the effect of early reionization due to global or local reheating (see e.g. Peebles, 1993, and White, Scott & Silk, 1994, for a review).

3 Instrumental and observational issues

3.1 Noise and systematics

There is a number of technical issues that make the detection of CMBR temperature fluctuations a difficult task. The first limitation comes from detector technology. Because of the very low level of these temperature fluctuations (a few tens of μK), extremely sensitive detectors are required. Two main technologies are currently available: phase sensitive radio detectors with state of the art high gain amplifiers using High Electron Mobility Transistors (HEMTs), and broad band receivers using cooled bolometers, sensitive to the total power of the radiation. For both technologies there a significant amount of noise, low frequency drifts and gain instabilities. Bolometer detectors tend to be more sensitive than HEMTs, especially at high frequencies (above 90 GHz), with the obvious drawback that they need to be cooled to sub-Kelvin temperature. In the case of ground-based and, to a less extent, of

balloon-borne experiments, atmospheric turbulence induces an important amount of additional noise. There is also a significant amount of atmospheric absorption in the frequency range of interest.

Two additional very important instrumental requirements are straylight rejection and thermal stability. Excellent straylight rejection is necessary since very strong radiation sources such as the earth, the moon and the sun, can induce signals that are typically of the same order of magnitude than the expected CMBR fluctuations, or even worse, when seen in the sidelobes of the radiation pattern of the antenna. Temperature fluctuations of all the elements of the instrument that are radiatively coupled to the detector (mirrors, feed horns, filters, radiation shields ...) also induce large variations in the signal. Stringent requirements have to be satisfied[4] which are almost impossible to achieve or even test for in terrestrial environment. Thus, it is very difficult, in the case of ground-based or balloon-borne experiments, to quantify exactly the effect of systematics on the anisotropy measurements.

3.2 Foregrounds

The next concern is that at the level of sensitivities required to recover interesting informations from primary and secondary anisotropies of the CMBR, it is expected that astrophysical pollution from so-called "foregrounds" may contribute significantly to the astrophysical signal. The main contributors to these foregrounds arise from the emission of our galaxy and are strongly concentrated towards the galactic plane. Three main processes contribute to the galactic emission in the interesting waveband between 10 and 1000 GHz : greybody emission from interstellar dust, synchrotron emission from electrons moving in the galactic magnetic field, and bremsstrahlung (or "free-free") emission of electrons off nuclei. To each of these processes corresponds an emission spectrum, which is known with some limited level of accuracy.

3.3 Component separation

Thus, several different astrophysical processes contribute to the total fluctuations detected. The primary CMBR fluctuations, the SZ effect, and the galactic foregrounds are the most important. Each of them has a particular frequency spectrum $S(\nu)$.

On most of the sky (away from the galactic plane and bright galaxy clusters), at frequencies around 100 GHz, the amplitude of primary CMBR fluctuations is expected to be, on average, an order of magnitude higher than that of fluctuations due to other processes. Nonetheless, there is much to be gained by evaluating as accurately as possible the contribution of each of the components: a good understanding of the systematics (necessary to test the statistics of CMBR fluctuations, and in particular to distinguish unambiguously between inflation-generated and defect-generated initial perturbations), and better accuracy on the measurement of CMBR anisotropies (and thus more accurate and reliable determination of cosmological parameters). Furthermore, each of the individual processes described above is by itself interesting for astrophysics or cosmology (in particular for the physics of galaxies and of clusters of galaxies, and for structure formation).

The simplest approach to component separation can be summarized in the following way. Consider n emission processes contributing to the total signal, having different frequency spectra of emission $S_1(\nu)$ ... $S_n(\nu)$. If these spectra are different, n different measurements of the total signal $S_{total}(\nu) = \alpha_1 S_1(\nu) + \alpha_2 S_2(\nu) + ... + \alpha_n S_n(\nu)$ at n well chosen frequencies ν_1, ν_2 ... ν_n allow to determine the constants α_i, at each point of the sky. Complications arise in the case of astrophysical component separation for several reasons. The first is that the spectra of some of the processes are not perfectly well known, or depend on several parameters. The second is that measurements available are not perfect measurements at a single frequency on individual pixels of the sky, but for each frequency channel j an integral over sky and frequency of a signal $S(\theta, \phi, \nu)$ multiplied by a channel-dependent transmission $T_j(\nu)$ and a beam pattern on the sky $A(\theta, \phi, \nu)$. Finally, noise and systematics are present in all the measurements.

[4] thermal stabilities in the range of 10 μK to a few mK, and sidelobe rejections towards earth and sun of 100 to 140 dB, depending on the instrument, the observing strategy...

This problem has been studied in some detail during the COBRAS/SAMBA phase A study. It has been demonstrated through numerical simulations that with an optimized set of frequency bands one could separate the components to an excellent level of accuracy over a large fraction of the sky, and obtain 10 arcmin resolution maps of primary CMBR fluctuations with a noise level of less than $6\mu K$ per pixel ($\Delta T/T < 2 \times 10^{-6}$) over more than one third of the sky (Bouchet *et al.*, 1996).

4 COBRAS/SAMBA

4.1 Mission concept

The COBRAS/SAMBA instruments have been optimized in order to cope with all the instrumental and observational problems described in the previous section. The orbit at the L2 sun-earth Lagrange point, 1.5 million kilometers away from the earth in the anti-sun direction, has been chosen so that systematic instrumental effects as sidelobe signals from the earth, sun and moon, and scan-synchronous fluctuations of the temperature of the mirrors and the focal plane, are minimized. The scan strategy allows post-mission recovery of the radiation pattern of the telescope, almost full-sky coverage, and adequate removal of drifts due to low-frequency noise. The large useful diameter of the primary mirror of the telescope allows for a good angular resolution in the bolometer channels, enabling measurements of the spectrum of CMBR temperature fluctuations for values of ℓ up to $\ell \sim 1500$.

Frequency channels have been chosen according to the results of the simulation of component separation. The highest two frequency channels allow a good measurement of the dust contribution, and the lowest two are dedicated to detecting the presence of significant Synchrotron and Free-Free. The 217 GHz channel is centered at the frequency where the SZ thermal effect vanishes, and the 353 GHz around the maximum of the positive thermal SZ spectral distortion. Outside of the galactic plane, the contribution of CMBR anisotropy is expected to dominate over other components between 60 and 200 GHz.

Frequency (GHz)	31.5	53	90	125	143	217	353	545	857
Type of detector	HEMT	HEMT	HEMT	HEMT	bolo	bolo	bolo	bolo	bolo
Resolution	30'	18'	12'	12'	10.3'	7.1'	4.4'	4.4'	4.4'
Bandwidth $\Delta\nu/\nu$	0.15	0.15	0.15	0.15	0.37	0.37	0.37	0.37	0.37
Sensitivity $\Delta T/T$ (10^{-6})	7.8	7.5	14.4	35.4	1.2	2.0	12.1	76.6	4166

Table 1: Summary of the COBRAS/SAMBA frequency channels, with resolution, bandwidth and sensitivity

It is important to stress that these characteristics are by far better than those of any other instrument, existing or planned, dedicated to CMBR anisotropy measurements. With COBRAS/SAMBA, accurate and reliable data will be made available to constrain the cosmological models.

All the details of the COBRAS/SAMBA mission concept can be found in the COBRAS/SAMBA report on the phase A study.

4.2 Scientific outcome

The COBRAS/SAMBA satellite will map the sky in 9 frequency channels between 30 and 850 GHz, with a resolution ranging from 30 to 4 arcminutes, depending on the channel. According to the results of the phase A study, this should allow:

1- to distinguish unambiguously between inflationary or defect-generated initial perturbations
2- to characterize these perturbations, and thus put constraints on very high energy physics
3- to put constraints on the nature of dark matter
4- to put constraints on scenarios of structure formation

5- to detect the SZ effect in thousands of clusters (Aghanim *et al.*, 1996)
6- if the Universe is at all as we imagine it is, to determine fundamental parameters (H_0, Ω_0, Ω_b, Λ...) to an accuracy better than a few percent
7- if not, to rule out models that would have been erroneously favored by cosmologists so far (although this option is not excluded, it is not very likely...), and provide a fantastic set of data to build better models
8- to enhance our understanding of the physics of our galaxy, of the properties of clusters of galaxies, of the infrared background of unresolved sources, and much more...

5 Conclusion

A lot of information about our Universe is still encrypted in the tiny fluctuations of the CMBR. The COBRAS/SAMBA mission, proposed by a european collaboration as a possible third medium-sized ESA mission, has been designed to constrain strongly the cosmological scenarios by measuring these fluctuations with unprecedented accuracy. If selected, it will revolutionize cosmology, since it will provide answers to a large number of open questions. Additionally, if current cosmological models are confirmed, it will provide accurate measurements of most cosmological parameters and constants.

REFERENCES

Aghanim, N., Désert, F.X., Puget, J.L., Gispert, R., 1995, Astronomy & Astrophysics, in press
Aghanim, N., *et al.*, 1996, in preparation
Albrecht, A., *et al.*, 1995, Phys. Rev. Lett., in press (astro-ph/9505030)
Bouchet, F.R, *et al.*, 1996, in preparation
Coulson, D., Ferreira, P., Graham, P., Turok, N., 1994, Nature 368:27
Crittenden, R.G., Turok, N.G., astro-ph/9509147
Efstathiou, G., *et al.*, 1996, in preparation
Fixsen, D. *et al.*, 1996, in press
Hu, W. & White, M., 1996, astro-ph/9602019, submitted to Ap.J.
Jungman, G., Kamionkowski, M., Kosowsky, A., Spergel, D. , 1996, submitted to Ap.J
Kamionkowski, M., Spergel, D.N, Sugiyama, N., 1994, Ap.J. 426:L57
Kolb, E.W., Turner, M.S., 1990, *The Early Universe*, Addison Wesley Publ.
Mather, J.C. *et al.*, 1993, Ap.J 432:L15
Smoot, G.F., Bennett, C.L, Kogut, A. *et al.*, 1991, Ap.J.Lett. 371:L1
Smoot, G.F., 1992, Ap.J. 396:L1
Smoot, G.F., Scott, D. , 1996, Astro-Ph/9603157
White, M., Scott, D., Silk, J., 1994, Ann. Rev. Astron. Astrophys. 32:319-370
Zel'dovich, Ya.B., Sunyaev, R.A., 1969, Astroph. Space Sci. 4:301

QUE PEUT-ON ENCORE APPRENDRE DU FOND DE RAYONNEMENT COSMOLOGIQUE ?

Les anisotropies du fond de rayonnement cosmologiques sont engendrées par toute une variété de processus physiques, depuis l'Univers Primordial dans la première seconde du Big-Bang, jusqu'à nos jours. Le satellite COBRAS/SAMBA, proposé par une collaboration européenne en réponse à l'appel d'offre pour la troisième mission moyenne de l'Agence Spatiale Européenne, a pour objectif la mesure de ces anisotropies avec une précision inégalée. Cette mesure permettrait de contraindre très fortement les modèles cosmologiques et, dans la plupart des cas de figure, de déterminer les valeurs des principaux paramètres cosmologiques comme H_0, Ω_0, ou Λ avec une précision de l'ordre du pourcent.

Measurement of q_0 with Type Ia Supernovae

R. Pain for the Supernova Cosmology Project[1]
LPNHE, CNRS-IN2P3 and Universités Paris VI & VII

Abstract : Our search for high-redshift Type Ia supernovae discovered, in its first years, a sample of seven supernovae. The spectra and light curves indicate that almost all were Type Ia supernovae at redshifts $z = 0.35$ to 0.5. These high-redshift supernovae can provide a distance indicator to measure the cosmological parameter q_0. In this paper, observation strategies, analysis and preliminary results on q_0, are presented.

Résumé : Dans une première étape de notre programme de recherche de supernovae à grand décalage vers le rouge, nous avons découvert 7 supernovae. Les spectres et courbes de lumière indiquent que pratiquement toutes sont de type Ia avec des décalages vers le rouge variant entre z=0.3 et z=0.5. Ces supernovae peuvent être utilisées comme indicateurs de distance pour mesurer le paramètre cosmologique q_0. Dans ce papier, les stratégies d'observation, l'analyse ainsi que les résultats préliminaires concernant la mesure de q_0 sont présentés.

1. Introduction

Since the mid 1980's, Type Ia supernovae (SNe Ia) have appeared likely to be homogeneous enough that they could be used for cosmological measurements. At the time, it appeared that they could be used to determine H_0, if their absolute magnitude at peak could be measured, i.e., if some SN Ia's distance could be calibrated. They could also be used to determine q_0 from the *apparent* magnitudes and redshifts of nearby and high redshift supernovae, if high redshift SNe Ia could be found. We will here discuss the search for high redshift supernovae and the measurement of q_0.

---

[1]S. Perlmutter, S. Deustua, S. Gabi, G. Goldhaber, D. Groom, I. Hook, A. Kim, M. Kim, J. Lee, R. Pain, C. Pennypacker and I. Small, Lawrence Berkeley National Laboratory and Center for Particle Astrophysics, University of California, Berkeley; A. Goobar, University of Stockholm; R. Ellis and R. McMahon, Institute of Astronomy, Cambridge University; B. Boyle, P. Bunclark, D. Carter, K. Glazebrook and M. Irwin, Royal Greenwich Observatory; H. Newberg, Fermi National Accelerator Laboratory; A.V. Filippenko and T. Matheson, University of California, Berkeley; M. Dopita and J. Mould, MSSSO, Australian National University; W. Couch, University of the new South Wales

2. High Redshift Supernovae Search Strategy

Type Ia Supernovae at high redshift are difficult to work with for at least three reasons: they are rare, they are rapid, and they are random. The estimates of SN Ia rates—a few per millennium per galaxy—are daunting, if one wants a statistically useful sample of supernovae. Much of the interesting data must be obtained rapidly, since the supernova rises to maximum light within a few weeks and, at high redshifts, fades below the largest telescopes' limits within a month or two. Furthermore, it is not possible to guarantee photometry and particularly spectroscopy of randomly occurring high-redshift supernovae, since the largest, most over-scheduled telescopes are needed to observe them.

To solve these problems, we developed a new search technique. Just after a new moon, we observe many tens of high-galactic-latitude fields (including known high-redshift clusters when possible) on a 2.5- to 4-meter telescope. With a wide-field camera, each image contains hundreds of galaxies at redshifts 0.3 – 0.6. Just before the following new moon, we observe the same fields again. We compare the images, thus checking tens of thousands of high redshift galaxies (including those below our detection limit) to find the ten or so showing the new light of a supernova that was not there on the previous observation. The supernovae generally do not have time to reach maximum light, with only 2.5 to 3 weeks (or approximately 11 to 14 days in the supernova rest frame) between our after- and before-new-moon comparison images. In order to begin the follow-up photometry and spectroscopy immediately, we have developed extensive software to make it possible to complete the analysis of all the images within hours of the observations.

We report here results concerning the first 9 supernovae found during 1994. Seven of these were found within the standard 2.5 – 3 week search interval and were followed with photometry and spectroscopy (Perlmutter *et al.* 1994, 1995a). (Since these were the demonstration runs of the project, not all of the follow up was scheduled.) We observed light curves for all of the supernovae in at least one filter (usually R band), and spectra for all of the host galaxies and three of the supernovae.

The results of the batch search strategy can be seen in Figure 1, which shows a preliminary analysis of the *R*-band light curves as of Dec 1995. For each of the supernovae, there is the 2.5 – 3 week gap in the observations leading up to the discovery just before maximum light. (There is also a gap during the following full moon.)

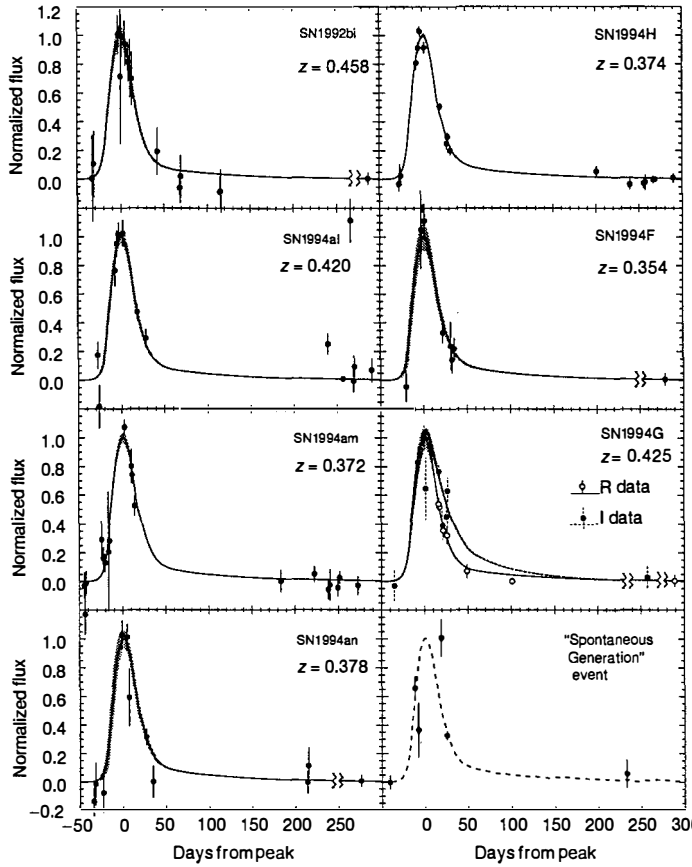


Fig. 1.— Preliminary R-band light curves for the first seven high-redshift supernovae. Shading behind the solid curves represent the 1-σ bounds on the best fit Leibundgut template light curve. An I-band light curve is also shown for SN 1994G; other photometry points in I and B for these supernovae are not shown on this plot. The lower of a "spontaneous generation" event, a transient point-source of light on a region of an image where no host galaxy is visible; the follow-up photometry was not as extensive for this event.

3. Photometry and Light Curve Analysis

The data analysis involves several stages, first to reduce the observed image data to individual photometry points and then to compare these points with nearby SN Ia light curves to determine the luminosity distance and hence q_0. The supernova light in each image must be measured and the underlying host galaxy light subtracted off.

To compare these high-redshift supernova photometry points to nearby SN Ia light curves, it is necessary to calculate the K correction that accounts for the redshifting of the light observed in a given filter. The standard K corrections give the magnitude difference between the light emitted in a given filter band and the light observed after redshifting in that same band. Since at redshifts of order $z = 0.45$ the light received in the R band corresponds approximately to the light emitted in the B band, we calculate a generalization of the K correction, K_{RB}, that gives the magnitude difference between the rest frame B magnitude and the observed R magnitude.

In the past year, the case has become quite strong that SNe Ia are a family of very similar events, not all identical. Hamuy *et al.* (1995) and Riess, Press, and Kirshner (1995) present evidence indicating that this family can be described by a single parameter, essentially representing the shape or width of the light curve, and that this parameter is tightly correlated with the absolute magnitude at maximum. The broad, slow-light-curve supernovae appear somewhat brighter, while the narrow, fast-light-curve supernovae are somewhat fainter.

Hamuy *et al.* used Phillips' (1993) characterizations of this light curve width, Δm_{15}, the magnitude drop in the first 15 days past maximum, and fit their data to template light curves from supernovae representing various Δm_{15} values. Riess *et al.* added and subtracted different amounts of a "correction template" to a Leibundgut "normal" template (Leibundgut *et al.* 1991 and references therein) to represent this same light curve variation.

For our preliminary analysis, we have chosen a third alternative approach, to stretch or compress the time axis of the Leibundgut template by a "stretch factor" s. We find that this simple parameterization gives variations on the light curve that fit the variety of SNe Ia and can be translated to Phillips' Δm_{15} via the formula $\Delta m_{15} = 1.7/s - 0.6$. Calibrating nearby supernovae with this s-factor, we find the width-brightness relation can be characterized by a magnitude correction of $\Delta\mathrm{mag} = 2.35(1 - s^{-1})$, which closely matches Hamuy *et al*'s Δm_{15} correction.

4. Preliminary Scientific Results

To use high-redshift SNe Ia as distance indicators, we must first ask if the width-brightness relation holds at $z \sim 0.4$. We plotted the stretch factor—or equivalently Δm_{15}—versus the K-corrected and Galactic-extinction-corrected absolute magnitude (for a given q_0 and H_0), and we found a correlation that is consistent with that found for nearby supernovae. Using the width-brightness relation as a magnitude "correction," established using only nearby SNe Ia, tightens the dispersion of the high redshift supernova K-corrected peak magnitudes from $\sigma_{\rm raw} = 0.32$ mag to $\sigma_{\rm corrected} = 0.21$ mag. These dispersions are quite comparable to those found for nearby SNe Ia before and after correcting for the width-brightness relation. For our preliminary estimate of q_0, we have applied the width-brightness correction to all seven of the high-redshift supernovae, on the assumption that they are all SNe Ia with negligible extinction and that the zero point of the width-brightness relation has not evolved in the ~4 billion years back to $z \sim 0.4$. Given the range of host galaxy ages for the nearby supernovae used to derive this relation, it is unlikely that it would have be very different at $z \sim 0.4$.

Figure 2 shows the results of this analysis plotted on a Hubble diagram, together with the relatively nearby SNe Ia of Hamuy *et al.* (1995). Comparing the upper and lower panels, it is clear that the scatter about the Hubble line decreases for both near and high-redshift supernovae after correcting for the width-brightness relation, although the error bars increase in the cases for which the light-curve width (s-factor or Δm_{15}) is poorly constrained by the photometry data. The three supernovae with the smallest error bars, after correction for the width-brightness relation, are all near the redshift $z = 0.37$ and their data points in Figure 2, lower panel, lie on top of each other. They favor a relatively high value for q_0. Note that before the width-brightness correction (Figure 2, upper panel) their data points spread from $q_0 \approx -0.5$ (upper dotted line) to $q_0 > 1$ (lower dotted line); generally, the correction has brightened as many points as it has dimmed.

5. Discussion and Conclusion

Given the error bars, our current measurements of q_0 do not yet clearly distinguish between an empty $q_0 = 0$ and closed $q_0 > 0.5$ universe. The data do, however, indicate that a decelerating $q_0 \geq 0$ is a better fit than an accelerating $q_0 < 0$ universe. This is an important conclusion since it limits the possibility that $q_0 = \Omega_0/2 - \Omega_\Lambda$ is dominated by the cosmological constant Λ (where Ω_Λ is the normalized cosmological constant $\Lambda(3H_0)^{-2}$). In an accelerating universe, high values for the Hubble constant do not conflict with the ages of the oldest stars, because the universe was expanding more slowly in the past. However,

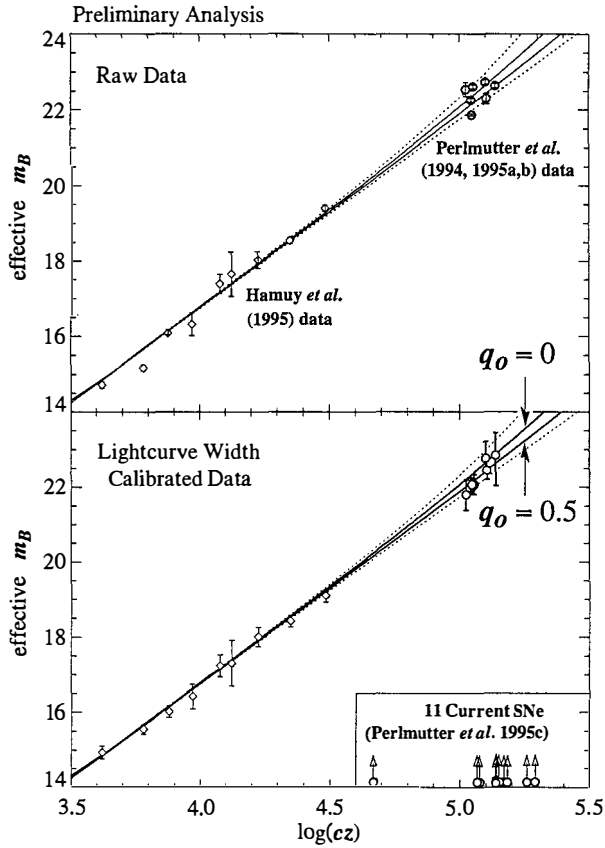


Fig. 2.— Hubble diagrams for first seven supernovae (preliminary analysis). Upper panel: "raw" apparent magnitude measurements, after K correction from R-band (observed) to B-band (supernova rest frame), versus redshift. The upper dotted curve is calculated for $q_0 = -0.5$, the two solid curves are for $q_0 = 0$ and $q_0 = 0.5$, and the lower dotted curve is for $q_0 = 1$. Lower panel: apparent magnitude measurements after "correction" for width-brightness relation. Note that the three points with the smallest error bars lie on top of each other, in this plot. Inset: Points represent redshifts for the most recent 11 supernova discoveries that are still being followed.

in a "flat" universe, in which $\Omega_0 + \Omega_\Lambda = 1$, a cosmological constant $\Omega_\Lambda \geq 0.5$ would give an acceleration $q_0 \leq -0.25$, a poor fit to our data.

Note that extinction of the distant supernovae would make this evidence of deceleration even stronger, as would the possibility of inhomogeneous matter distribution in the universe as described by Kantowski, Vaughan, & Branch (1995), since both of these would lead to underestimates of q_0 if not taken into account.

It will be important to compare the q_0 value from our most distant supernovae to the value from our closest high-redshift supernovae to look for Malmquist bias. An estimate of the size of this bias using our studies of our detection efficiency as a function of magnitude suggests that a 0.2 mag intrinsic dispersion in calibrated SN Ia magnitudes would lead to an overestimate of 0.1 in q_0, if not accounted for.

We have shown here that the scheduled discovery and follow up of batches of pre-maximum high-redshift supernovae can be accomplished routinely, that the comparison with nearby supernovae can be accomplished with a generalized K-correction, and that a width-brightness calibration can be applied to standardize the magnitudes. Ideally this will now become a standard method in the field, and SNe Ia beyond $z = 0.35$ will become a well-studied distance indicator—as they already are for $z < 0.1$—useful for measuring the cosmological parameters. The prospects for this look good: already several other supernova groups have now started high-redshift searches.

The observations described in this paper were primarily obtained as visiting/guest astronomers at the Isaac Newton and William Herschel Telescopes, operated by the Royal Greenwich Observatory at the Spanish Observatorio del Roque de los Muchachos of the Instituto de Astrofisica de Canarias; the Kitt Peak National Observatory 4-meter and 2.1-meter telescopes and Cerro Tololo Interamerican Observatory 4-meter telescope, both operated by the National Optical Astronomy Observatory under contract to the National Science Foundation; the Keck Ten-meter Telescope; and the Siding Springs 2.3-meter Telescope of the Australian National University. We thank the staff of these observatories for their excellent support. Other observers contributed to this data as well; in particular, we thank Marc Postman, Tod Lauer, William Oegerle, and John Hoessel for their more extended participation in the observing. This work was supported in part by the Physics Division, E. O. Lawrence Berkeley National Laboratory of the U. S. Department of Energy under Contract No. DE-AC03-76SF000098, and by the National Science Foundation's Center for Particle Astrophysics, University of California, Berkeley under grant No. ADT-88909616. The author would like to thank Gerard Fontaine of CNRS-IN2P3 and Bernard Sadoulet of CfPA, Berkeley for encouraging his participation to the project.

REFERENCES

M. Hamuy, M. M. Phillips, J. Maza, N. B. Suntzeff, R. A. Schommer, and R. Aviles. *Astron. J.*, 109:1, 1995.

R. Kantowski, T.E. Vaughan, and D. Branch. *Ap. J.*, 447:35, 1995.

A. Kim, A. Goobar, and S. Perlmutter. *Pub. Astr. Soc. Pacific*, 1996, in press.

B. Leibundgut, G.A. Tammann, R. Cadonau, and D. Cerrito. *Astro. Astrophys. Suppl. Ser.*, 89:537, 1991.

P. Nugent, M. Phillips, E. Baron, D. Branch, and P. Hauschildt. *Ap.J.Lett.*, 1996, in press.

S. Perlmutter et al. *Ap.J.*, 440:L41, 1995a.

S. Perlmutter et al. International Astronomical Union Circular, 1994, nos. 5956 and 5958.

S. Perlmutter et al. International Astronomical Union Circular, 1995b, no. 6263.

S. Perlmutter et al. International Astronomical Union Circular, 1995b, no. 6270.

M. M. Phillips. *Ap.J.*, 413:L105, 1993.

A. G. Riess, W. H. Press, and R. P. Kirshner. *Ap.J.*, 438:L17, 1995.

T. Vaughan, D. Branch, D. Miller, and S. Perlmutter. *Ap.J.*, 439:558, 1995.

ON PHYSICALLY SELFCONSISTENT BASIS FOR LARGE SCALE STRUCTURE FORMATION THEORY

Maxim Yu. Khlopov
Center for CosmoParticle Physics "Cosmion"
125047, Moscow, Russia, e-mail khlopov@khlopov.rc.ac.ru

Abstract

The modern Theory of Large Scale Structure Formation contains several elements, which may be related to physical mechanisms of inflation, baryosynthesis and physics of dark matter. These mechanisms are predicted by particle theory, but have no experimentally proven grounds. The cosmological framework, spectrum of primordial fluctutions and transfer function, defined by the properties of dark matter particles, baryonic/dark matter density ratio and relative distribution have different physical grounds or motivations and generally are treated independently. Physically selfconsistent treatment corresponds to unique theoretical framework for their description. Such a description will provide both quantitatively definite basis for LSS formation scenarios and the possibilities for its detailed test by the combination of additional effects, predicted in the respective theoretical framework. The example of the cosmological scenarios, based on horizontal unification is considered.

The modern theory of large scale structure implies both cosmological framework and mechanisms of structure formation as necessary elements. The widely accepted viewpoint is to ascribe the observed structure of inhomogeneities to the result of gravitational instability in the big bang Universe. So one uses 1) cosmological parameters, 2) spectrum of primordial fluctutions, 3) dark matter content, defining the transfer function between the density fluctutions and 4) baryonic/dark matter density ratio and relative distribution as independent input parameters (See for review [1] and Refs. wherein) for calculations and numerical simulations to confront the observational data on the cosmological large scale structure. Cosmological tradition is to treat independently the framework, i.e. the big bang scenario, and the physical processes in it, the mechanisms of structure formation, in particular. So a wide variety of structure formation models, assuming different values of Ω , Λ, dark matter and baryonic content, follow in fact this tradition. But the proper choice between them, as well the possible role and effect of more sophisticated mechanisms of structure formation, such as textures, cosmic strings, late phase transitions etc can not be done in the framework of cosmology only. However exact the data on the large scale structure would be, they are insufficient to make definite conclusion on the origin of the structure, accounting for all the reasonable mechanisms offered by the particle theory. The aim of the present note is to stress, that cosmoparticle physics[2], studying the foundations of cosmology and particle physics in their fundamental relationship sheds new light onto the problems of large scale structure formation theory and opens the way to its physically selfconsistent treatment.

The close relationship between cosmology and particle physics in the problems of large scale structure (LSS) formation is widely recognized. One needs the existence of nonbaryonic dark matter to make consistent the formation of the observed large scale structure from small initial fluctuations with the observed level of isotropy of thermal electromagnetic background radiation, putting constraint on the amplitude of such fluctuations. From purely cosmological viewpoint it corresponds to additional parameter(s) in the theory of large scale structure formation. In the simpliest case of a single form of stable dark matter species, whithin the frame of inflational cosmology, putting, in general, the total density very close to the critical one, and assuming "flat" Harrison-Zeldovich (HZ) spectrum one additional parameter is involved, being in the most popular models, either the free-streaming scale of HDM, or biasing in relative baryonic and dark matter distribution in CDM. One may invoke two-parameter mixed HDM+CDM or unstable dark matter models, or consider some cosmological consequences of particle theory leading to cosmic string, texture, late phase transition etc mechanisms of structure formation. The other direction of extensions is to relax the inflational condition of critical density, to include nonzero cosmological Λ term or to relax the condition of flat spectum and/or Gaussian nature of initial density fluctuations. In most of these trends new parameters are included on the base of their own motivation, without any possible correspondance with the other parameters of the LSS formation theory. The choice between the models is strongly correlated in this case to the matter of taste, to their simplicity or popularity.

However, the modern cosmological paradygm, treating the big bang Universe as inflational model with baryosynthesis and nonbaryonic dark matter opens the way to physically motivated cosmological scenarios, in which both the cosmological framework and parameters of physical processes in it follow from the respective physical grounds. In this approach flattenness of the modern Universe reflects the existence of inflational stage[3], the observed baryon asymmetry is determined by the process of baryosynthesis[4], the large scale structure formation and the modern total density are provided by the existence of nonbaryonic dark matter[1,5]. All these phenomena follow from the particle theory, so that the cosmological parameters Ω, B/γ, $\delta\rho/\rho$, Ω_{dm} may be treated selfconsistently on its base. The problem is, that the physical grounds for inflation, baryosynthesis and dark matter are related to hypothetical phenomena, predicted by particle theory but having no direct experimental proofs. Moreover, cosmological consistency is in its turn an important test for these predictions of particle theory. The mutual relationship of the foundations of

cosmology and particle physics, so that the resolution of their internal problems can not be separated, is the bed-rock of cosmoparticle physics, chasing the solution of this puzzle in the complex of indirect physical, astrophysical and cosmological means, probing simultaneously the basics of both cosmology and particle theory.

There are two complementary cosmoparticle approaches which may be useful on the way to physically selfconsistent theory of LSS formation. Methods of cosmoarcheology[6], treating the astrophysical data as the experimental sample in the Gedanken Experiment-al study of cosmological consequences of particle theory, establish the correlation in these consequences specific for the particular physical mechanism, underlying inflation, baryosynthesis and related with the physical nature of the dark matter. One may combine then the predictions of the considered model for LSS with additional effects, predicted on its base, so that different physical mechanisms may be distiguished by their different effects on chemical composition and spatial distribution of baryonic matter , in spectrum and angular distribution of blackbody electromagnetic background radiation, in nonthermal electromagnetic (rf, IK, optical, UV, X-ray and γ), neutrino and gravitational wave backgrounds. In the case of LSS formation it provides both the analysis of the physical basis for the cosmological framework and for the mechanisms of LSS formation. It means, that the choice of particular physical mechanism for inflation, baryosynthesis and of particular candidate for dark matter leads to specific influence on the spectrum of density fluctuations and/or initial inhomogeneities and on the mechanisms of structure formation, correlated with additional effects in the phenomena, lying generally outside the scopes of the LSS formation theory.

In the modern cosmological paradygm inflation[3] explains the global properties of the modern Universe, its flattenness and similarity of conditions in apparently causally disconnected regions. So the main cosmological parameter of LSS theory, Ω, (as a rule, Ω=1) finds physical grounds in inflational scenarios. In many cases it also provides the mechanism for almost flat Harrison-Zeldovich type spectrum for initial density perturbations. But in practically all the realizations of inflational scenario post-inflational dust-like stage is to take place after inflation before the cosmological reheating. Such a stage may be sufficiently long to provide[7] the formation of primordial black holes (PBH) [8], and the astrophysical constraints on the PBH spectrum [9] put additional restrictions on LSS formation theory parameters. The data, constraining the presence of PBHs in the modern Universe or the effects of their evaporation, also put restrictions on the multicomponent inflation, in which non-flat spectrum or peaks in it, reflecting phase transitions at inflational stage, may appear (See [10] for review).

Baryosynthesis is considered in the modern cosmology as the physical reason for baryon asymmetric Universe. It provides the physical origin to the important cosmological parameter B/γ. However, practically all the mechanisms of baryosynthesis may, under some conditions, be mechanisms of anti-baryo-synthesis, since the value and even the sign of baryon excess, generated in baryogenesis, may vary in space. In the original Sakharov's scenario[4], where baryon excess is the result nonequilibrium CP violating baryon non conserving processes, spatial variation of the baryon excess B(x) reflects the spatial variation of CP violating phase ϕ(x), arising in some mechanisms of CP violation (See for rev. [6,11,12]). In SUSY GUT induced scenarios baryon excess is ascribed to primordial condesate of scalar quarks, for which spatial dependence of B(x) is possible. Under the condition of spontaneous CP violation spatial variation of baryon excess also may appear in the mechanisms of baryon excess generation due to electroweak nonconservation of B at high temperatures[13]. It makes possible the existence of antimatter domains even in the baryon asymmetric Universe, so that search for cosmic antimatter and/or for effects of its annihilation in the early Universe[11,12] turns to be interesting probe for the physical grounds for baryon asymmetric cosmology and not only for nontrivial realizations of baryon symmetrical Universe[14]. The principal possibility for spatial variation of baryon excess seem to be the common feature of all the

mechanisms of baryosynthesis, thus adding, in general, to LSS formation theory spectrum of entropy $B(x)/\gamma$ fluctuations as additional input parameter.

Dark matter explains the LSS formation at the observed level of isotropy of the black body background radiation. In the simpliest case its physical realization is the gas of relic weakly interacting particles with nonzero rest mass. In the widely assumed picture (See rev in [1]) dark matter provides the growth of initial fluctuations and its parameters define the dissipational scale in the transfer function relating the spectrum of primordial fluctutions with the one in the beginning of matter dominancy stage. In the simpliest case of dark matter particles, frozen out or decoupled from the thermal equilibrium in the early Universe, this dissipational scale, determining the minimal size of inhomogeneities, formed by the dark matter, is defined by the mass of particles. So, the simpliest realization of Hot Dark Matter (HDM) corresponds to few (ten)eV mass of neutrinos and Cold Dark Matter (CDM) to few (ten) eV neutrallinos. However, physical grounds even for these simpliest cases may lead to additional nontrivial cosmologically significant effects. The see-saw mechanism of neutrino mass generation invokes heavy right-handed Majorana neutrino, which may be sufficiently stable to dominate in the early Universe and form PBHs at the stage of their dominancy[8,15]. The mechanisms for supersymmetry breaking may predict the existence of metastable gravitinos, making decaying gravitinos nontrivial element of physical processes in the early Universe [16]. The more complicated dark matter content and its underlying physics is, the more additional effects may be predicted in cosmology. Moreover the physical motivations favor in general multicomponent dark matter models[17].

Cosmoarcheology traces these effects and arranges cosmoarcheological chains, linking the considered hypothetical phenomenon to respective astrophysical data in order to test its consistency. The definite answer here is negative, since each trace, taken separately can definitely put the constraint only: the predicted effect should not contradict the observational data. To find definitely positive answer, corresponding to the data, the physically selfconsistent treatment of all the necessary cosmological elements should take place.

Such approach may be illustrated by the recently proposed model of horizontal unification[18]. The extension of the standard $SU(2)xU(1)xSU(3)_c$ model of electroweak and strong (QCD) interactions to the gauge symmetry $SU(3)_H$ of quark and lepton families provides not only reasonable theoretical description of the established existence of three families of quarks and leptons (ν_e,e,u,d); (ν_μ,μ,c,s); (ν_τ,τ,t,b), but in its realization[19] turns to be the theoretical framework, incorporating in an unique scheme physical grounds for inflation (in the role of inflaton - singlet scalar field, used in quark and lepton mass generation by Dirac see-saw mechanism,), baryosynthesis (by the combination of nonequilibrium CP violating lepton number violating processes an electroweak baryon nonconservation at high temperatures) and dark matter candidates (HDM -massive neutrinos, CDM - invisible axions, sharing the properies of Majorons and familons and called archions and UDM[20] - heavier neutrinos, decaying on lighter neutrinos and archions).

Even in the present form the model[18] offers the quantitatively definite correspondance between fundamental cosmological parameters (form of inflaton potential, lepton number violation, mass, spectrum and lifetime of dark matter particles and fields), astrophysical effects (rate of stellar archion emittion, contributing significantly stellar energy losses and dynamics of stellar collapse) and particle properties (see-saw mechanism of mass generation, hierachy of masses and mixings of quark and lepton families, Majorana mass ratio of neutrinos, rates of archion decays, double neutrinoless beta decays). Finally, the amount of free parameters of the model[18] turns to be much less, than the amount of its signatures in particle processes, astrophysics and cosmology, thus providing its definite test.

On reproducing the observed masses and mixings of quarks and leptons the main free parameter of the model[18] is the *a priori* unknown scale of family symmetry breaking, F. Complex analysis of the set of its physical, astrophysical and cosmological predictions makes it possible to fix

the value of F in two rather narrow windows (around 10^6 and around 10^{10} GeV). The both solutions correspond to LSS formation theory with flat HZ spectrum of density fluctuations and differ in the dark matter content. In the first model, corresponding to $F=10^6$ GeV, LSS is formed in the succession of stages of dominancy of massive neutrinos and their decay products, combining the advantages of HDM, CDM and UDM models. The second solution, corresponding to $F=10^{10}$ GeV, seem to reproduce all the main features of widely assumed as standard cosmological scenario with inflation, baryosynthesis and cold (axionic) dark matter. However, even this most simple reduced cosmological scenario does contain some additional elements (cf post inflational dust-like stage, on which PBH formation is possible, effects of successive PBH evaporation, formation of primordial percollational structure of archioles etc). This example favors the conclusion, that in *no* cases cosmological effects of of realistic particle theory are reduced to inflation, baryosynthesis and dark matter *only*, making physically consistent LSS formation theory nontrivial mixture of different existing approaches. It also resembles the system of nontrivial crossdisciplinary links which should be used to probe the true physically consistent cosmology and, in particular, LSS formation theory with the use of the methods of cosmoparticle physics.

The work was undertaken in the framework of the project "Cosmoparticle physics" and International project "Astrodamus". I express my gratitude to J.Tran Thanh Van and Rencontres de Moriond for hospitality.

References

1. K.Olive, This Volume. M.Lachieze-Rey, This Volume. F.Bernardeau. This Volume.

2. A.D.Sakharov. Cosmoparticle physics - crossdisciplinary problem. Vestnik AN SSSR, 1989, N-4, 39. M.Yu.Khlopov. Cosmoparticle physics. Ser. Fizika. Znanie 89/3. M. 1989. 140. M.Yu.Khlopov. Cosmoparticle physics. The way to study the foundations of cosmology and particle physics. in: Cosmoparticle physics 1, Proc. 1 International conference on cosmoparticle physics "Cosmion-94", dedicated to 80 Anniversary of Ya.B.Zeldovich and 5 Memorial of A.D.Sakharov, Moscow, December 5-14, 1994. Eds. M.Yu.Khlopov, M.E.Prokhorov, A.A.Starobinsky, J.Tran Thanh Van, Editions Frontieres, 1996. pp.1-8

3. A..Guth. Phys.Rev. D23,347-356, 1981. See review in: Inflationary Universe. Eds. L.F.Abbott, S.-Y. Pi.World Scientific. 1986

4. A.D.Sakharov. JETP Lett.5, 17-20, 1967. V.A.Kuzmin. JETP Lett., 12, 335-337, 1970.

5. See for review G.Jungman. This Volume

6. M.Yu.Khlopov. Cosmoarcheology. Direct and indirect astrophysical effects of hypothetical particles and fields. in: Cosmoparticle physics 1, Proc. 1 International conference on cosmoparticle physics "Cosmion-94", dedicated to 80 Anniversary of Ya.B.Zeldovich and 5 Memorial of A.D.Sakharov, Moscow, December 5-14, 1994.Eds. M.Yu.Khlopov, M.E.Prokhorov, A.A.Starobinsky, J.Tran Thanh Van, Editions Frontieres, 1996. pp.67-76.

7. M.Yu.Khlopov, B.A.Malomed, Ya.B.Zeldovich. Gravitational instability of scalar field and primordial black holes. Mon.Not.Roy.astr.Soc., 214, 100-114, 1985.

8.Khlopov M.Yu.,Polnarev A.G.. Primordial black holes as a test of GUT. Phys.Lett.,B97,383-387,1980 A.G.Polnarev, M.Yu.Khlopov . Stages of superheravy particle domination in the early Universeand primordial black holes. Astron.Zh., 58, 706-715, 1980.

9.Polnarev A.G.,Khlopov M.Yu.. Cosmology, primordial black holes and superheavy particles. Usp. Phys. Nauk. ,145,369-398,1985 [Sov.Phys. Usp.,28,213-227,1985] Khlopov M.Yu.,Polnarev A.G..Superheavy particles in cosmology and evolution of inhomogenities in the early Universe.in: The very early Universe. Eds.G.W.Gibbons,S.W.Hawking,S.T.C.Siclos. Cambridge.1983,Repr.1984,papbk 1985.PP.407-440. M.Yu.Khlopov, A.G.Polnarev, P.D.Naselsky Primordial black holes and observational test of quantum gravity. Proc. III Int. Sem. "Quantum gravity", 1985. With

10. A.S.Sakharov, M.Yu.Khlopov. Cosmological signatures of family symmetry breaking in multicomponent inflation models. Yad.Fiz., 56, 220-230, 1993 [Sov.J.Nucl.Phys., 56, 412-417, 1993].

11.Khlopov M.Yu.,Chechetkin V.M.. Antiprotons in the Universe as a cosmological test of grand unification. Fiz. Elem. Chastits At.Yadra,18,627,1987 [Sov.J. Part.Nucl.,18,267-288,1987]. V.M.Chechetkin, M.Yu.Khlopov, M.G.Sapozhnikov, Ya.B.Zeldovich,Astrophysical aspects of antiproton interaction with He (Antimatter in the Universe). Phys. Lett., 118B, 329-333, 1982. Chechetkin V.M.,Khlopov M.Yu., Sapozhnikov M.G.. Antiproton interactions with light elements as a test of GUT cosmologies.Riv.Nuovo Cimento, 5, N-10, 1982. Khlopov M.Yu.. Astronuclear experiment ASTROBELIX. Znanie, ser. Fizika, 1990/7,1990; Astrophysical implications of antiproton nuclear interactions in astronuclear experiment Astrobelix. Sov. J. Nucl.Phys., 55,765-766,1992

12. Astrodamus collaboration. Astroparticle studies of the physical nature of dark, decaying and annihilating matter in the Universe. in: Cosmoparticle physics 1, Proc. 1 International conference on cosmoparticle physics "Cosmion-94", dedicated to 80 Anniversary of Ya.B.Zeldovich and 5 Memorial of A.D.Sakharov, Moscow, December 5-14, 1994. Eds. M.Yu.Khlopov, M.E.Prokhorov, A.A.Starobinsky, J.Tran Thanh Van, Editions Frontieres, 1996. pp.99-106. Towards cosmoarcheology of cosmic rays and gamma backgrounds of the Universe. Ibid, pp.107-112
13. Comelli D., Pietroni M., Riotto A.. Spontaneous CP violation and baryogenesis in the minimal supersymmetric standard model.Nucl.Phys.B412, 441-458, 1994
14. A.DeRujula. This Volume.
15. Berezhiani Z.G.,Khlopov M.Yu.. Physical and astrophysical effects of fermion families symmetry breaking. Yad.Fiz.,51,1479- 1491,1990 [Sov.J.Nucl.Phys.,51,1990]
16. F.Balestra, I.V.Falomkin, G.Piragino, M.G.Sapozhnikov, D.B.Pontecorvo, M.Yu.Khlopov. Antiproton-helium-4 annihilation at low energies and its relationship with the problems of modern cosmology and GUT models.Yad.Fiz., 39, 990-997, 1984. [Sov.J.Nucl.Phys.,39, 626-630, 1984]. M.Yu.Khlopov, A.D.Linde. Is it easy to save gravitino? Phys. Lett., 138B, 265-268, 1984.
17. M.Yu.Khlopov . Physical arguments, favouring multicomponent dark matter. in. Dark matter in cosmology, clocks and tests of fundamental laws. Eds.B.Guiderdoni, G.Greene, D.Hinds, J.Tran Thanh Van. Editions Frontieres, 1995.PP 133-138
18. Sakharov A.S.,Khlopov M.Yu.. Horizontal unification as phenomenology of "Theories of Everthing". Yad. Fiz., 57,690-697,1994. M.Yu.Khlopov, A.S.Sakharov. Cosmoparticle physics of family symmetry breaking. In: Particle astrophysics, atomic physics and gravitation. Eds. J.Tran Thanh Van, G.Fontaine, E.Hinds. Editions Frontieres, 1994. pp. 197-207.
19. .Berezhiani Z.G.,Khlopov M.Yu.. Cosmology of spontaneously broken gauge family symmetry with axion solution of strong CP-problem. Z.Phys.C -Particles and fields,49,73-78,1991
20. .Doroshkevich A.G.,Khlopov M.Yu..Formation of the structure in the Universe with unstable neutrinos. Mon.not.Roy.astr.Soc.,211, 277-283,1984; Doroshkevich A.G.,Khlopov M.Yu.,Klypin A.A..Structure formation by unstable dark matter particles. Mon.not.Roy.astr.Soc., 239,923-930,1989.

QUANTUM MEASUREMENT

A PRECISION TEST OF DECOHERENCE

J.R. Anglin and W.H. Zurek
Theoretical Astrophysics, T-6, Mail Stop B288
Los Alamos National Laboratory
Los Alamos, New Mexico 87545

ABSTRACT: The motion of a charged particle over a conducting plate is damped by Ohmic resistance to image currents. This interaction between the particle and the plate must also produce decoherence, which can be detected by examining interference patterns made by diffracted particle beams which have passed over the plate. Because the current densities within the plate decay rapidly with the height of the particle beam above it, the strength of decoherence should be adjustable across a wide range, allowing one to probe the full range of quantum through classical behaviour.

Introduction

Charged particles passing close to conducting surfaces have been investigated in tests of the Ahoronov-Bohm effect [1]) and the equivalence principle [2]). The phenomenon of dissipation due to electrical resistance to image currents in the conductors has been investigated extensively, but the effects on the particles of the concomitant resistor noise have been discounted [3]). Among these effects will be decoherence[4]). We propose that the strong dependence of the image current densities to the easily varied parameter of trajectory height makes the motion of a charged particle over a conducting plate an excellent precision test for our understanding of decoherence.

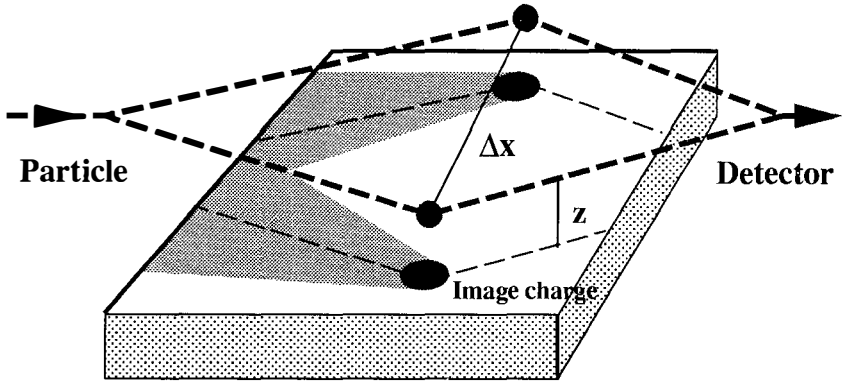


Figure 1: Sketch of proposed system. The heavy dashed lines indicate two trajectories of the particle over the conducting plate. The large shaded regions represent the disturbance in the electron gas inside the plate.

Decoherence

If a charged particle passes over a conducting plate along some trajectory, it induces an image charge on the surface of the conductor, which moves with the particle by inducing bulk currents within the body of the conductor [5]). These bulk currents encounter Ohmic resistance, which is primarily due in metals at room temperature to the scattering of coherent electron flow by thermal phonons. The passage of the charged particle along a trajectory over the conductor thus leaves a wake of disturbance in the electron and phonon gas beneath it.

After the charge has crossed the plate (and continued on its flight towards a detector), the disturbed state of the conductor remains as a record of the charge's passage (and in fact, of the particular path that it took). Quantum mechanically, this means that the state of the electron gas and of the charged particle must be described together by a joint wave function. The contribution to this wave function from a single trajectory $\mathcal{T}$, at the end of which the particle is at position Q, can be written as

$$|\Psi_{\mathcal{T}}\rangle = |Q\rangle\,|\psi_E[\mathcal{T}]\rangle\ .$$

According to Feynman's picture of quantum evolution, the final state of the system is found by a weighted sum of such states $|\Psi_{\mathcal{T}}\rangle$ over all trajectories $\mathcal{T}$ which the particle may have taken. Quantum mechanics then predicts the probability of finally observing the particle at a given position Q (without measuring any properties of the conducting plate) to be

$$P(Q) = \sum_{Q_i,Q_i'} \phi(Q_i)\phi^*(Q_i') \times \sum_{\mathcal{T}}\sum_{\mathcal{T}'} e^{\frac{i}{\hbar}(S[\mathcal{T}]-S[\mathcal{T}'])}\langle\psi_E[\mathcal{T}']|\psi_E[\mathcal{T}]\rangle\ ,$$

where the sums are over trajectories $\mathcal{T}$ or $\mathcal{T}'$ in which the particle begins at Q_i or Q_i' and ends up at Q, and $\phi(Q_i)$ is the wave function describing the initial state of the particle.

If this initial state is a superposition of two distinct localized states, such as is produced by a double slit, cross terms in (99) can produce a probability distribution $P(Q)$ which is not the classical "two lumps", but an interference pattern characteristic of wave mechanics. The familiar textbook case in which this possibility is fully realized occurs only when $\langle\psi_E[\mathcal{T}']|\psi_E[\mathcal{T}]\rangle \rightarrow 1$, so that the conducting plate does not effectively distinguish between different particle trajectories. In general, though, this inner product is *not* unity, and trajectories that are far apart may disturb the plate's electron gas in such different places that the inner product of the electron states will be negligible. Interference between these trajectories will therefore be suppressed, and the final probability distribution of the particle will thus be altered towards the classical limit.

This phenomenon is an example of the process known as *decoherence*[4]), which is believed to play a crucial role in enforcing the effective classicality of macroscopic physics, and is thought to be the greatest challenge facing such hypothetical advanced technologies as quantum computing[6]). A precision test of decoherence is thus highly desirable.

Precision Test of Decoherence

The system of charged particle and conducting plate is suitable for a precision test of decoherence, because the strength of their interaction is adjustable over a wide range. Calculations using classical electromagnetism, for a charge Q moving over the plate at a height z and constant velocity v show

a rate of Joule heating in the conductor which is

$$P_C = \frac{Q^2 \rho v^2}{16\pi z^3} ,$$

where ρ is the specific resistivity of the plate. This implies an Ohmic dissipation rate proportional to d^{-3}: we can compute the relaxation time τ_r to be

$$\tau_r = \frac{v}{\dot{v}} \sim \left(\frac{z}{10^{-4}\mathrm{m}}\right)^3 \times \begin{cases} 2 \times 10^3 \mathrm{s} & \text{electron} \\ 3\frac{M}{Q^2} \times 10^6 \mathrm{s} & \text{ion} \end{cases}$$

where M is the ion mass in units of proton mass, and we take the resistivity ρ to be on the order of $10^{-6}\Omega$m. (This rather high resistivity is possessed by manganese at room temperature; since a resistivity a few orders of magnitude higher still could be even more convenient, alloys or semi-conductors might be contemplated instead of a pure metal.)

(Note that adding a thin layer of insulator on top of the conducting plate will actually produce dissipation proportional to the inverse *fourth* power of z [2,5]). The sensitivity to z that we are considering can thus be made even greater; but we will not discuss this option here in detail.)

These calculations are classical, and a proper quantum treatment is certainly required to make serious theoretical predictions concerning decoherence in this system. We must consider the unitary evolution of phonons and a nearly free electron gas as a charge passes over the plate on an arbitrary trajectory. The electromagnetic field may presumably be treated classically, and for sufficiently slow particles an adiabatic expansion to first order in the particle speed will probably suffice. After the charge has crossed the plate, the direct perturbation of the conductor by the charge will cease, but the resistive interactions among the conductor's constituents that were driven by the charge, as it dragged an image current of electrons through the mill of

thermal nuclei, will have brought the conductor into a final quantum state that depends on the particle's trajectory. The inner product between two such states must then be computed.

For the present brief communication, we merely suppose that the full quantum calculation will recover both the classical dissipation rate, and the result (typical in linear systems) that decoherence rates are proportional to it. As a zeroth order approximation that should be reasonable at room temperature, we will assume that the decoherence time scale τ_d is given by the formula for a completely linear model at high temperature[7]),

$$\tau_d = \tau_r \Big(\frac{\lambda_{dB}}{\Delta x} \Big)^2 ,$$

where $\lambda_{dB} \equiv h/\sqrt{2mk_BT}$ is the thermal de Broglie wavelength, and Δx is a length scale characterizing the difference between the two quantum states that are to decohere. In our case, we will take this scale to be the distance between Feynman trajectories across the plate, which could perhaps be of order 10^{-4}m.

Since τ_r is proportional to the particle mass, we therefore estimate the same decoherence time for singly charged ions and electrons:

$$\tau_d \sim \Big(\frac{z}{10^{-4}\mathrm{m}} \Big)^3 \times 10^{-5}\mathrm{s} .$$

If we assume that our charged particles will traverse a conducting plate that measures a centimeter across, flying at a speed on the order of a kilometers per second, the decoherence time will be on the order of the time of flight for $z \sim 0.1$ mm. By varying the trajectory height z we should therefore indeed be able to probe the full range from negligible to strong decoherence.

Conclusion

There are a number of aspects of decoherence which one would like to have clarified by a precision experiment. One of the most interesting is the dependence of the decoherence of two trajectories on the spatial distance between them. In simple linear models, decoherence has a Gaussian profile with distance, but in general we expect that this profile will instead approach some non-zero constant value at large distances [8,9]). In the particle and conductor system, we expect the length scale at which this saturation occurs to be set by the correlation length in the electron gas.

By varying the separation between the two slits through which the particle is coherently passed before crossing the plate (or using different crystal lattices for diffraction, if a low intensity electron beam is used instead of a stream of ions), one can vary the distances over which quantum interference of trajectories is important in determining the final $P(Q)$. This should allow one to probe the dependence of decoherence on spatial separation. By also varying the overall strength of decoherence, through the trajectory height d, we can investigate decoherence in a two dimensional parameter space, and provide a stringent test of decoherence in a system which is theoretically non-trivial. Unlike set-ups deliberately designed to mimic simple theoretical models for decoherence, the particle crossing a conducting plate will thus be able to stretch our theory: it is a real precision test of decoherence.

References

[1] T.H. Boyer, Phys. Rev. D **8**, 1667 (1973); **8**, 1679 (1973).

[2] T.W. Darling *et al.*, Rev. Mod. Phys. **64**, 237 (1992).

[3] L.H. Ford, Phys. Rev. D **47**, 5571 (1993).

[4] W.H. Zurek, Phys. Rev. D **24**, 1516 (1981); *Physics Today*, October 1991.

[5] T.H. Boyer, Phys. Rev. A **9**, 68 (1974).

[6] I.L. Chuang, R. Laflamme, P. Shor, and W.H. Zurek, *Science* **270**, 1633 (1995).

[7] W.H. Zurek, in G.T. Moore and M.O. Scully, eds., *Frontiers of Nonequilibrium Statistical Physics* (Plenum: New York, 1986).

[8] M.R. Gallis and G.N. Fleming, Phys. Rev. A **42**, 38 (1990); **43**, 5778 (1991).

[9] J.R. Anglin, J.P. Paz, and W.H. Zurek, "Deconstructing Decoherence", in preparation.

French title and abstract:

UN TEST PRÉCIS DE DÉCOHÉRENCE

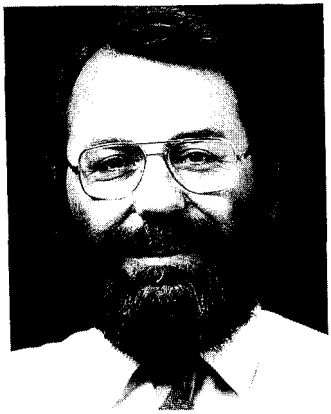
Le mouvement d'une particule chargée au dessus d'un un conducteur plat est dissipée par la résistance Ohmique due aux courants images. Cette interaction entre la particule et le conducteur produit aussi de la décohérence, qu'on peut détecter dans les franges d'interférence formées par des faisceaux de particules qui ont traversé le conducteur. Puisque la densité de courant dans le conducteur est très sensible à la hauteur de la particule, on peut controller le degrée de décohérence, et à cette façon étudier le phénomène de décohérence du domaine quantique jusqu'à celui classique.

Propensities: Measuring Quantum States with Quantum Rulers

P. L. Knight

Optics Section, Blackett Laboratory, Imperial College,

London SW7 2BZ, England



1 Introduction

In this paper I describe recent work in quantum optics on the measurement of quantum fluctuations. A single field mode of frequency ω is quantized in a box of volume V, and described by the field operator [1]

$$\hat{E} = (\frac{2\omega^2}{\epsilon_0 V})^{\frac{1}{2}} \hat{q} \sin kz \tag{1}$$

where ϵ_0 is the permittivity of free space and $k = \omega/c$ the wavevector; $\hat{q}$ is an Hermitian operator having the dimension of a length. This can be expressed in terms of annihilation ($\hat{a}$) and creation operators ($\hat{a}^\dagger$) as

$$\hat{E} = \sqrt{2}\mathcal{E}_0 \sin kz \{\hat{a}\, e^{-i\omega t} + \hat{a}^\dagger e^{i\omega t}\} \tag{2}$$

where $\hat{a} = (2\hbar\omega)^{-\frac{1}{2}}(\omega\hat{q} + i\hat{p})$, $\quad [\hat{a}, \hat{a}^\dagger] = 1$, where $\hat{p}$ is a conjugate field variable to $\hat{q}$. The quantity $\mathcal{E}_0 = (\hbar\omega/2\epsilon_0 V)^{\frac{1}{2}}$ is the "electric field per photon'. Alternatively, we can parametrize

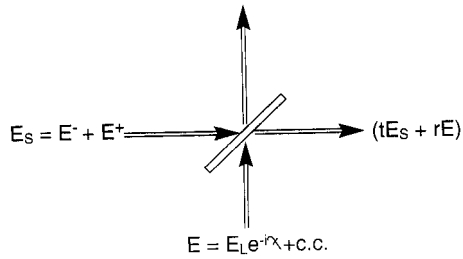


Figure 1: Homodyne mixing of a system field of interest E_s with a local field of amplitude E_l and adjustable phase χ by a beam splitter of transmissivity t and reflectivity r.

the field in terms of the quadrature operators $\hat{X}$ and $\hat{Y}$ so that $\hat{E} = \sqrt{2}\mathcal{E}_0 \sin kz \{\hat{X}\cos\omega t + \hat{Y}\sin\omega t\}$, where $\hat{X} = (1/2)(\hat{a}+\hat{a}^\dagger)$ and $\hat{Y} = (1/2)(\hat{a}-\hat{a}^\dagger)$ with $[\hat{X},\hat{Y}] = i/2$ and $\Delta X \Delta Y \geq 1/4$. The variables X and Y describe the phase–space of the field mode. One can interfere the light field of interest with a reference beam from a well–stabilized local oscillator, in homodyne phase–sensitive detection.

The field $(tE_s + rE_l)$ shown in Fig. 1 is detected by a square law device responding to the intensity. The detected signal thus contains an interference term $I_i \sim \langle E_s E_l^* e^{i\chi} + c.c.\rangle$. If the local oscillator is prepared in a coherent state this can be factored out of the expectation value as $I_i \sim E_l(\langle E^{(-)}\rangle e^{-i\chi} + \langle E^{(+)}\rangle e^{+i\chi}) = E_l\langle E(\chi)\rangle$. Normally $rE_l \gg t\langle E\rangle$ so that the detected intensity $\langle I_d\rangle$ is given by [2] $\langle I_d\rangle \cong |r|^2|E_l|^2 + |t||r||E_l|\langle E(\chi)\rangle$. A suitable choice of phase angle χ allows us to read off the value of the system quadrature X and Y directly. The photon number fluctuations are given by

$$\langle(\Delta n)^2\rangle_{det} \sim |r|^2|E_l|^2\{|r|^2 + 4|t|^2\langle(\Delta E(\chi))^2\rangle\}. \tag{3}$$

If the quantum noise on the system–field is phase dependent then as χ is varied, the number fluctuations also vary and for some phase angles these may be reduced below that of the normal coherent state "shot noise" result. We can write $\omega t = \chi$, to describe the general phase-dependence of the electric field through [2],

$$\hat{E}(\chi) = \sqrt{2}\mathcal{E}_0 \sin kz\,(\hat{a}\,e^{-i\chi} + \hat{a}^\dagger e^{i\chi}) \tag{4}$$

where the phase χ may be due to the insertion of optical elements, in an experimental situation to retard the field. The total electric field operators evaluated at different phase-angles do not commute:

$$[\hat{E}(0), \hat{E}(\chi)] = i\,4\mathcal{E}_0^2 \sin\chi \tag{5}$$

so that the associated field uncertainties $\Delta E(0)$ and $\Delta E(\chi)$ satisfy the uncertainty relation

$$\Delta E(0)\,\Delta E(\chi) \geq 2\mathcal{E}_0^2\,|\sin\chi|\,. \tag{6}$$

For most field states, the uncertainty is *independent* of χ. But if the uncertainty at one phase (say $\chi = 0$) is manipulated to be small (compared with $\mathcal{E}_0$), then at a later phase angle the uncertainty is correspondingly increased: this is the basis of *squeezed light* [2].

2 Phase–Space and Quasi–Probabilities

A classical stochastic model of fluctuating light fields centers on the joint probability density $P(q,p;t)$ for finding the field with quadratures X, Y or equivalently of "position" q and "momentum" p in phase space. This joint probability density is positive everywhere, and is normalized. A classical joint probability $P(q,p;t) \equiv P(\alpha,\alpha^*;t)$ where α is some complex amplitude, can be related to a characteristic function $C(\beta,\beta^*;t)$ through the Fourier transform

$$C(\beta,\beta^*;t) = \int d^2\alpha\, \exp(\beta\alpha^* - \beta^*\alpha)\, P(\alpha,\alpha^*,t). \tag{7}$$

Cahill and Glauber [3] defined the "s–ordered" *quantum* characteristic function $C(\beta,s)$ by

$$C(\beta,\beta^*;s) = Tr\{\exp(\beta\hat{a}^\dagger - \beta^*\hat{a})\rho\}\exp(-|\beta|^2\frac{s}{2}). \tag{8}$$

If $s = 0$ this generates the Wigner function W, whereas if $s = -1$ then we obtain the Glauber–Sudarshan diagonal P–function and if $s = +1$ then we obtain the Husimi, or Q–function. For the Wigner function, the mean value of some observable $\hat{M}$ is given by

$$\langle M\rangle = Tr\{\hat{\rho}M\} = \frac{1}{\pi}\int d^2\beta W(\beta)M(\beta). \tag{9}$$

The coherent states $|\alpha\rangle$ are generated by the Glauber displacement operator $\hat{D}(\alpha) = \exp(\alpha\hat{a}^\dagger - \alpha^*\hat{a})$ so that $|\alpha\rangle = \hat{D}(\alpha)|0\rangle$. The observable $\hat{M} = \hat{M}(\hat{q},\hat{p})$ has as its phase–space counterpart $M(q,p) = M(\beta)$. The Wigner function can be written as

$$W(\beta) = \frac{1}{\pi}\int C^W(\beta)\exp(\beta\eta^* - \beta^*\eta)\,d^2\eta \tag{10}$$

where $C^W(\beta)$ is the Weyl–ordered (symmetric) characteristic function

$$C^W(\eta) \equiv Tr[\hat{\rho}\hat{D}(\eta)] \tag{11}$$

so for example,

$$\langle\{\hat{a}^\dagger\hat{a}\}^W\rangle = \langle\frac{1}{2}(\hat{a}^\dagger\hat{a} + \hat{a}\hat{a}^\dagger)\rangle = \frac{1}{\pi}\int d^2\beta\, W(\beta). \tag{12}$$

The marginals of the Wigner function for an oscillator with $\hat{a} = (2\hbar)^{-1/2}(\lambda\hat{q} + i\lambda^{-1}\hat{p}$ (where $\lambda = (m\omega)^{1/2}$ for a material oscillator of mass m) are

$$W(q) = \frac{1}{\lambda(2\pi\hbar)^{\frac{1}{2}}}\int dp\, W(q,p) = \frac{(2\pi\hbar)^{\frac{1}{2}}}{\lambda}\langle q|\hat{\rho}|q\rangle \tag{13}$$

and moments can be calculated from

$$\langle\hat{q}^n\rangle = \frac{\lambda}{(2\pi\hbar)^{\frac{1}{2}}}\int dq\, q^n\, W(q). \tag{14}$$

For coherent states, $\langle(\Delta q)^2\rangle = \langle\hat{q}^2\rangle - \langle\hat{q}\rangle^2 = \frac{\hbar}{2\lambda^2} = \hbar\sigma_q^2$ and $\langle(\Delta p)^2\rangle = \langle\hat{p}^2\rangle - \langle\hat{p}\rangle^2 = \frac{\hbar\lambda^2}{2} = \hbar\sigma_p^2$ so that $\langle(\Delta q)^2\rangle\langle(\Delta p)^2\rangle = \frac{\hbar^2}{4}$. Wigner functions in general cannot be positive everywhere. Indeed, the only Wigner function which is positive everywhere [4] is the Gaussian, *i.e.* the coherent states and the M.U.S. squeezed states.

3 Joint Measurements of p, q in Phase Space

Consider a beam–splitter used to generate two output modes, such that one may be used to measure q and one to measure p. The difficulties with this kind of joint measurement scheme stems from the unused input port: vacuum noise injected from this port introduces excess noise of the kind considered many years ago by Arthurs and Kelly [5]. What permits the simultaneous measurement of (p, q) is in fact the unused port: the output mode variables actually commute as they represent system *plus* meter variables; because of the unused port, we can read out the q-variable of one output and p-variable of the other output mode. These can therefore be read separately without adding extra noise beyond that represented by the beam–splitter which superposes the system and the meter variables to generate the two output modes.

It is helpful to compare the system state with that of a reference state (sometimes referred to as a quantum ruler) [6], and to employ Popper's concept of "propensity" [7]. The propensity P_r describes the tendency of the measured object described by the system's density operator $\hat{\rho}_s$ to take up certain states prescribed by a measuring device (or filter) described by a projection

operator $P_{\Phi_f} = |\Phi_f\rangle\langle\Phi_f|$. In comparing system and meter states, one normally employs the probability $Tr(\hat{\rho}_s P_{\Phi_f})$. But imagine tuning the filter, or quantum ruler in phase space: the filter projection operator can be tuned by the action of some unitary transformation U_g by varying some parameter g in the filter:

$$P_{\Phi_f} \longrightarrow U_g^\dagger|\Phi_f\rangle\langle\Phi_f|U_g = P_{\Phi_f}(g) \tag{15}$$

For example, the filter states can be displaced in phase space by the operator $U(q,p) = exp(iq\hat{p}/\hbar + i\hat{q}p/\hbar)$. The propensity can then be written as $P_r = \text{const.}Tr(\hat{\rho}_s P_{\Phi_f}) = \text{const.}$ $Tr(\hat{\rho}_s \hat{U}_g^\dagger|\Phi_f\rangle\langle\Phi_f|\hat{U}_g)$ so that we may probe $\hat{\rho}_s$ by moving the filter around to sample all attributes 'g'. For the case of displacements,

$$P_r(q,p) = \frac{1}{2\pi\hbar} Tr(\hat{\rho}_s \hat{U}_g^\dagger(q,p)|\Phi_f\rangle\langle\Phi_f|\hat{U}_g(q,p)) \tag{16}$$

and if the system is prepared in the pure state $\hat{\rho} = |\Psi_f\rangle\langle\Psi_f|$, then

$$P_r(q,p) = \left|\int dx e^{-ipx/\hbar}\Phi_f^*(q+x)\Psi_s(x)\right|^2 \tag{17}$$

$$= \frac{1}{(2\pi\hbar)}\int dqdp W_\Psi(q+q', p+p')W_\Phi(q',p') \tag{18}$$

In other words the propensity is given by the convolution of system and quantum ruler Wigner functions. In terms of the Glauber displacement operator $\hat{D}(q,p)$, this may be written as

$$P_{\Psi\Phi} = P_r(q,p) = |\langle\Psi|\hat{D}(q,p)|\Phi\rangle|^2 \tag{19}$$

which shows clearly that the propensity is always positive. The ruler states act as a smoothing in this convolution process.

Now let us return to the problem of joint measurement of q and p. If the quantum ruler is prepared in the vacuum state $|\Phi\rangle = |0\rangle$

$$P_{\Psi\Phi} = |\langle 0|\hat{D}^\dagger(q,p)\hat{\rho}_\Psi \hat{D}(q,p)|0\rangle|^2 = Q_\Psi(q,p) \tag{20}$$

where $Q_\Psi(q,p)$ is the Husimi Q–function. One way of realizing this operational phase–space measurement is to use the 8–port homodyne device shown in Fig. 2 in which a beam–splitter generates two output fields, each of which is used in a balanced homodyne measurement, one for q, the other for p. Of course what is measured is the system coordinates *plus* the meter (or quantum ruler) noise.

In ordinary homodyne detection, the fields and photon number counts can be easily found [2] from the usual beam–splitter transformation from inputs $\hat{\boldsymbol{a}}_L$ (for local oscillator) and $\hat{a}$ to output modes $\hat{d}_1, \hat{d}_2$ as

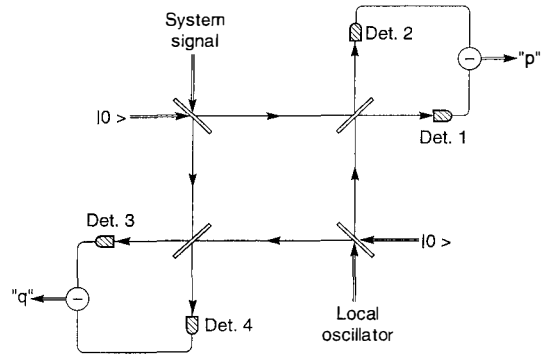


Figure 2: 8–port homodyne set–up for joint measurements in phase–space

$$\begin{pmatrix} \hat{d}_1 \\ \hat{d}_2 \end{pmatrix} = \begin{pmatrix} r & t \\ t & r \end{pmatrix} \begin{pmatrix} \hat{a}_L \\ \hat{a} \end{pmatrix}$$

where the amplitude reflection and transmission coefficients are denoted by r and t. Unitarity demands that

$$|r|^2 + |t|^2 = 1, \tag{21}$$

$$t^* r + r^* t = 0. \tag{22}$$

In ordinary homodyne detection, $|r|^2 \ll |t|^2$, but in balanced homodyne detection $|r| = |t| = 1/\sqrt{2}$. In ordinary homodyne detection the the measured signal is determined by the operator.

$$\hat{d}_1^\dagger \hat{d}_1 = |r|^2 \hat{a}_L^\dagger \hat{a}_L + t^* r \hat{a}_L^\dagger \hat{a}_L + r^* t \hat{a}_L^\dagger \hat{a}_L + |t|^2 \hat{a}^\dagger \hat{a}. \tag{23}$$

If the local oscillator field is prepared in coherent state $|\alpha\rangle_L$, then,

$$\langle \hat{n}_1 \rangle = \langle \hat{d}_1^\dagger \hat{d}_1 \rangle = |r|^2 |\alpha_L|^2 + 2|t||r||\alpha_L| \langle E(\chi) \rangle + |t|^2 \langle \hat{a}^\dagger \hat{a} \rangle \tag{24}$$

where$\chi = \arg r - \arg t + \boldsymbol{\phi}_L = \frac{\pi}{2} + \phi_L$. In ordinary homodyne detection, the strong local oscillator allows us to write the detected photon number approximately as

$$\langle \hat{n}_1 \rangle \cong |r|^2 |\alpha_L|^2 + 2|t||r||\alpha_L| \langle E(\chi) \rangle \tag{25}$$

and the photon number variance as

$$\langle (\Delta n_1)^2 \rangle = |r|^2 |\alpha_L|^2 [|r|^2 + 4|t|^2 | \langle (\Delta E(\chi))^2 \rangle] \tag{26}$$

In practice the photodetectors have a finite quantum efficiency η. This can be modelled [8] by considering the finite efficiency detector to be equivalent to a perfect photodetector preceded by a beam–splitter which transmits only a fraction $\sqrt{\eta}$ of the incident amplitude. The annihilation operator at the detector can be written as

$$\hat{d} = \sqrt{\eta}a + \sqrt{1-\eta}a_v \tag{27}$$

where a_v represents the unavoidable admixture of vacuum field from the other input to the beam–splitter. The field represented by $\hat{d}$ is then assumed to be detected with unit efficiency. Then $\langle n \rangle = \eta\langle \hat{a}^\dagger \hat{a} \rangle$ and $\langle \hat{d}^\dagger \hat{d}^\dagger \hat{d}\hat{d} \rangle = \eta^2 \langle \hat{a}^\dagger \hat{a}^\dagger \hat{a}\hat{a} \rangle = \eta^2 \langle \hat{n}(\hat{n}-1) \rangle$. The *photocount* mean and variance are $\langle \hat{m} \rangle = \eta \langle \hat{n} \rangle$,
$\langle (\Delta m)^2 \rangle = \eta^2 \langle (\Delta n)^2 \rangle + \eta(1-\eta)\langle \hat{n} \rangle$. Here the '$n$"s are really the photon flux in quanta per unit time multiplied by the integration time T. So for the ordinary homodyne detection process, we can write (after subtracting the large local oscillator contribution) the mean photocount as

$$\langle \hat{m} \rangle = 2\eta |t||r||\alpha_L| \langle E(\chi) \rangle \tag{28}$$

and its variance as

$$\langle (\Delta m)^2 \rangle \cong \eta |r|^2 |\alpha_L|^2 \{1 + \eta |t|^2 [4\langle (\Delta E(\chi))^2 \rangle - 1]\}. \tag{29}$$

The first "1" term in the above is the shot noise contribution, and the remaining term vanishes for a coherent signal input. If the input signal is squeezed, with $\langle (\Delta E(\chi))^2 \rangle < 1/4$ for some phase angle χ, then the photocount variance is reduced *below* the shot–noise limit.

For *balanced* homodyne detection, the difference in the two photodetector counts is measured: the relevant quantity is $\hat{d}_1^\dagger \hat{d}_1 - \hat{d}_2^\dagger \hat{d}_2 = i(\hat{a}^\dagger \hat{a}_l - \hat{a}_l^\dagger \hat{a})$ and the mean difference photon number is

$$\langle n_{12} \rangle = \langle \hat{d}_1^\dagger \hat{d}_1 - \hat{d}_2^\dagger \hat{d}_2 \rangle = 2|\alpha_L| \langle E(\chi) \rangle \tag{30}$$

with variance (again assuming $|\alpha_L|$ is large)

$$\langle (\Delta n_{12})^2 \rangle = 4|\alpha_L|^2 \langle (\Delta E(\chi))^2 \rangle. \tag{31}$$

The balancing removes the local oscillator contribution to the noise. For the *photocounts*, taking into account the finite efficiency we find

$$\langle m_{12} \rangle = 2\eta |\alpha_L| \langle E(\chi) \rangle \tag{32}$$

$$\langle (\Delta m_{12})^2 \rangle = \eta |\alpha_L|^2 \{1 + \eta [4\langle (\Delta E(\chi))^2 \rangle - 1]\}. \tag{33}$$

If we now use these (4-port) homodyne results to describe the 8–port joint measurement scheme, then we find for the difference photocounts in the 1–2 and 3–4 detectors $\langle m_{12}\rangle = \eta|\alpha_L|\langle E(\chi + \frac{\pi}{2})\rangle\langle m_{34}\rangle = \eta|\alpha_L|\langle E(\chi)\rangle$ so that the 8–port device measures the two quadrature components of the input field, with phase determined by χ. A suitable adjustment of this allows us to measure X (or q) in the 1–2 pair, and Y in the 3–4 pair. The photocount difference variances are [2]

$$\langle(\Delta m_{12})^2\rangle = \frac{1}{4}\eta|\alpha_L|^2\{2 + \eta[4\langle(\Delta E(\chi + \frac{\pi}{2}))^2\rangle - 1]\} \tag{34}$$

$$\langle(\Delta m_{34})^2\rangle = \frac{1}{4}\eta|\alpha_L|^2\{2 + \eta[4\langle(\Delta E(\chi))^2\rangle - 1]\} \tag{35}$$

If we compare these 8–port variances with the 4–port result eq.(47), we see the "1" factor in the curly brackets has been replaced by "2". This is because the photocount variances are no longer determined solely by the input ((system) variances alone (even in the case of perfect detectors). The additional contributions to the photocount fluctuations in the 8–port device can be traced back to the entrance of vacuum fluctuations in the two unused ports. The additional noise [2] is precisely the Arthurs–Kelly excess noise which takes us from the Wigner function of the input to the smoother, broader, Husimi Q–function measured in the joint–detection of the conjugate variables. The experimental realization of this 8–port homodyne joint measurement of q, p was undertaken by Walker and Carroll [9]. As far as I know, such measurements have not been attempted for nonclassical fields.

If our system is prepared in a coherent state $|\psi\rangle = |\alpha\rangle$, this is a minimum uncertainty state with $\langle(\Delta\hat{q})^2\rangle_\psi = \langle\psi|(\Delta\hat{q})^2|\psi\rangle = \hbar/2 = \langle(\Delta\hat{p})^2\rangle$. We can define the joint *marginals* from the propensities through

$$P_{\psi\phi}(q) = \frac{1}{\lambda\sqrt{(2\pi\hbar)}}\int dp P_{\psi\phi}(p,q) \tag{36}$$

$$P_{\psi\phi}(p) = \frac{\lambda}{\sqrt{(2\pi\hbar)}}\int dq P_{\psi\phi}(p,q). \tag{37}$$

These carry information about system *and* ruler. For example, consider the expectation value of the position operator

$$\langle\hat{q}\rangle_{\psi\phi} = \frac{\lambda}{\sqrt{(2\pi\hbar)}}\int dq q P_{\psi\phi}(q) \tag{38}$$

$$= \langle\psi|\hat{q}|\psi\rangle - \langle\phi|\hat{q}|\phi\rangle. \tag{39}$$

Similarly, for the momentum operator we find $\langle\hat{p}\rangle_{\psi\phi} = \langle\psi|\hat{p}|\psi\rangle - \langle\phi|\hat{p}|\phi\rangle$. These averages give the *relative* "distance" of the detected state with respect to the reference fixed by the quantum

ruler state. We can therefore define the *operational* variances :

$$\Sigma^2_{\psi\phi}(q) = \langle \hat{q}^2 \rangle_{\psi\phi} - \langle \hat{q} \rangle^2_{\psi\phi} \quad (40)$$

$$= \langle (\Delta \hat{q})^2 \rangle_{\psi} + \langle (\Delta \hat{q})^2 \rangle_{\phi} \quad (41)$$

$$\Sigma^2_{\psi\phi}(p) = \langle (\Delta \hat{p})^2 \rangle_{\psi} + \langle (\Delta \hat{p})^2 \rangle_{\phi} \quad (42)$$

which directly exhibit the *two* sources of measured quantum fluctuations: those from the system *and* those from the filter. The product of the two is:

$$\Sigma^2_{\psi\phi}(q)\Sigma^2_{\psi\phi}(p) \geq \hbar^2 \quad (43)$$

which is a factor of 4 greater than those of the system alone. For the simple case of the system prepared in a coherent state, $\Sigma^2_{\psi\phi}(q) = \hbar = \Sigma^2_{\psi\phi}(p)$ and the product is $\hbar^2$, exhibiting Arthurs–Kelly excess noise even for these coherent field inputs. We have used the propensity concept to discuss operational phase–space entropies [10]: again the uncertainties in system and quantum ruler can be distinguished. Further developements will, I hope, involve joint measurements of nonclassical field states.

Acknowledgements

I thank V. Bužek, C.H. Keitel,R. Loudon, M.B. Plenio, and V. Vedral for many discussions on the subject of this paper, and V. Vedral for help with the manuscript. This work was supported in part by the UK Engineering and Physical Sciences Research Council and the European Union.

References

[1] M. Sargent III, M.O. Scully and W.E. Lamb Jr., "Laser Physics",(Addison-Wesley, Reading, 1975).

[2] R. Loudon and P.L. Knight, J. Mod. Opt. **34**, 209 (1987).

[3] K.E. Cahill and R.J Glauber, Phys. Rev. **177**, 1857 (1969); ibid **177**, 1882 (1969)

[4] R.L. Hudson, Reports on Math. Phys. **6**, 249 (1974); N. Lütkenhaus and S.M Barnett, Phys. Rev. **A 51**, 3340 (1995).

[5] E. Arthurs and J.L. Kelly, Bell. Syst. Tech. J. **44**, 725 (1965).

[6] Y. Aharanov, D.Z. Albert and C.K. Au, Phys. Rev. Lett. **47**, 1029 (1981); R.F O'Connell and A.K. Rajagopal, Phys. Rev. Lett. **48**, 525 (1982); K. Wodkiewicz, Phys. Rev. Lett. **52**, 1064 (1984); K. Wodkiewicz, Phys. Lett. **115**, 304 (1986).

[7] K. Popper, "Quantum Theory and the Schism in Physics", (Hutchinson, London, 1982), p. 125–130.

[8] H.P. Yuen and J.H. Shapiro, IEEE Trans. Inf. Theory **26**, 78 (1980).

[9] N.G. Walker and J.E. Carroll, Electron. Lett. **20**, 981 (1987); N.G. Walker and J.E. Carroll, Opt. Qu. Electron. **18**, 355 (1986); N.G. Walker, J. Mod Opt. **34**, 15 (1987).

[10] V. Bužek, C. Keitel and P.L. Knight, Phys. Rev. **A 51**, 2575 and 2594 (1995).

Résumé

Cet article décrit comment le bruit quantique contribue aux processus de détection en optique quantique, et comment il est représenté dans l'espace des phases. Les mesures simultanées des variables canoniques sont déterminées non seulement par les propriétés du système, mais également par celles de la "règle quantique" (ou cosystème utilisé pour realiser la mesure).

Titre: "Propensities": Mesure d'états quantiques à l'aide de "règles quantiques"

QUANTUM MEASUREMENT AND REGENERATION OF A SINGLE SUPERPOSITION STATE

Y. Yamamoto, I. Chuang, O. Alter, and J. M. Jacobson
ERATO Quantum Fluctuation Project
Edward L. Ginzton Laboratory, Stanford University, Stanford, CA 94305

Introduction

Quantum correlation and nonlocality of a superposition state features various amusing phenomena and, in the future, may provide powerful means for secured privacy for communications and massive-parallel information processing.[1,2] Unfortunately, such a superposition state is a fragile object and easily destroyed by dissipative coupling to external worlds. This is especially true when two or more states are macroscopically separated: the effect is known as the "Schrödinger cat paradox."

A key concept for providing robustness in classical communications and information processing systems is "regeneration." Regeneration prevents a signal from being heavily distorted due to attenuation, dephasing, additive noise, and imperfections, and suppresses multiplicative increases in the bit error rate. Translation of this concept into the quantum domain is not straightforward because complete information for a single superposition state cannot be extracted by any measurement scheme. One obvious solution to this problem is to introduce redundancy into the system, i.e., many quantum bits (qubits) are used to represent one bit of information and a majority vote is taken (a technique known as error correcting codes).[3,4]

In this paper we will discuss an alternative approach to preserve qubit information.[5] The new scheme is based on the representation of one bit of information by dual-rail qubits and a balanced quantum nondemolition measurement. We begin with two application examples of superposition states, i.e., the photonic de Broglie wave interferometer[6] and a quantum computer solving the Deutsch problem.[7] We then discuss the impossibility of measuring a

single superposition state,[8] and finally describe a quantum regeneration scheme.

Photonic de Broglie Wave Interferometer

The time evolution of a photon number eigenstate $|n\rangle$ is described by the unitary operator $\hat{U} = exp(-i\omega\hat{n}t)$ which is obtained by standard procedure using the free-field Hamiltonian $\hat{\mathcal{H}} = \hbar\omega\left(\hat{n} + \frac{1}{2}\right)$. Therefore, the state vector evolves with angular frequency ωn instead of ω. This means that such a photon number eigenstate $|n\rangle$ has a "collective" momentum $P = \hbar\left(\frac{\omega}{c}\right) n$ and, thus, an associated de Broglie wavelength of $\lambda_{dB} = \frac{\hbar}{P} = \frac{\lambda}{n}$ can be assigned to photon number state $|n\rangle$, where λ is an ordinary optical wavelength.

In order to measure the photonic de Broglie wavelength λ_{dB}, an interferometer must be constructed in which photon number eigenstate $|n\rangle$ simultaneously propagates in two arms as a (macroscopically separated) superposition state and acquires a phase difference corresponding to the photonic de Broglie wavelength, i.e., $|\psi\rangle = \frac{1}{\sqrt{2}}\left[|n\rangle_1|0\rangle_2 e^{i\omega n\ell_1} + |0\rangle_1|n\rangle_2 e^{i\omega n\ell_2}\right]$, where ℓ_1 and ℓ_2 are the upper and lower arm lengths of the interferometer.

A conventional 50-50% beamsplitter cannot achieve this task because it splits each individual constituent photon into the two arms (not n photons, as a whole). A high-Q cavity containing a single atom prepared in a superposition state, $|\phi\rangle = \frac{1}{\sqrt{2}}(|g\rangle + |e\rangle)$, can be used for this purpose.[9]. The high-Q cavity simultaneously features 100% reflectivity and 100% transmission, similar to an atom sitting simultaneously in the ground state and in the excited state. The interferometer setup is shown schematically in Fig. 1.[6]. The first two $\frac{\pi}{2}$ pulses driving the single atom prepare the two cavities in the form of "intelligent" beamsplitters and the third $\frac{\pi}{2}$ pulse eliminates the information on which path the photon number state $|n\rangle$ actually takes. The atomic state (either $|e\rangle$ or $|g\rangle$) must be measured after the third $\frac{\pi}{2}$ pulse and activate the two detectors only when the measurement result is $|e\rangle$.

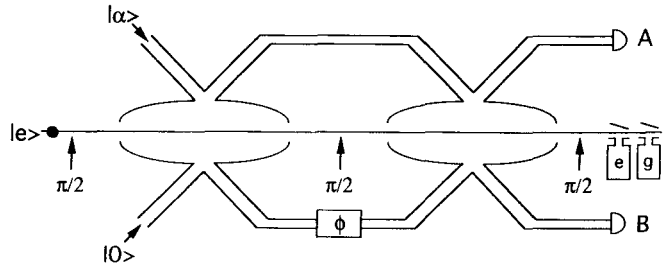


Fig. 1. A photonic de Broglie wave interferometer.[6]

An alternative way to construct a photonic de Broglie wave interferometer is to use a quantum optical Fredkin gate with two inputs $|n\rangle$ and $|0\rangle$ and one control signal prepared in a superposition state $|\phi\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$, where $|0\rangle$ and $|1\rangle$ are the vacuum and single-photon states, respectively.[10]

Figure 2 shows the average counts at one detector vs. the arm length difference when a coherent state $|\alpha\rangle$, with an average photon number $\bar{n} = 100$, is incident on the apparatus shown in Fig. 1. The oscillatory behavior with the photonic de Broglie wavelength $\frac{\lambda}{\bar{n}}$ is clearly seen. A coherent state consists of many different photon number states and, thus, slightly different photonic de Broglie wavelengths of constituent photon number states smear out the interference in the large path length difference.

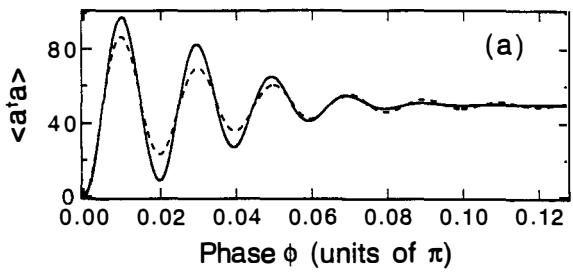


Fig. 2. Average photon count vs. phaseshift for a coherent state $|\alpha = \sqrt{100}\rangle$.[6]

It is easily shown that the minimum detectable phaseshift $\Delta\phi_{min}$ in the photonic de Broglie wave interferometer, defined by the signal-to-(quantum) noise ratio $S/N = 1$, is given by $\Delta\phi_{min} = \frac{1}{n}$, where n is the photon number per measured time interval.[6] This sensitivity is known as the fundamental quantum limit and is superior to the standard shot noise limit $\Delta\phi_{min} = \frac{1}{\sqrt{n}}$ because n is usually much greater than one.

The potential applications of a photonic de Broglie wave have not been thoroughly studied. A photon number eigenstate $|n = 10^3\rangle$ at an optical wavelength $\lambda = 5000\mathring{A}$, for instance, can feature the interference period of $\frac{\lambda}{n} = 5\mathring{A}$, a fact that is very attractive for applications to optical lithography and microscopy with x-ray resolution using visible light.

Quantum Computer Solving the Deutsch Problem

The Deutsch problem is the following game:[11] Alice sends Bob one integer x out of $[0, 2L-1]$. Bob calculates a function $f(x)$ which is either one or zero and sends it back to Alice. Bob's function is either Type I [$f(x)$ is all 0's or all 1's] or Type II [half of $f(x)$ is zero

and the other half is 1]. How fast can Alice determine the type of Bob's function? This is not a useful problem for practical applications but a quite pedagogical example demonstrating the power of "quantum parallelism." A classical solution, based on Alice's $N(= log_2 2L)$ bit query and Bob's one-bit reply, obviously requires $L+1(= 2^{N-1}+1)$ query and reply sessions. On the other hand, Alice can obtain a correct answer only by two query and reply sessions if she can use a quantum superposition state which simultaneously represents all possible $x \in [0, 2L-1]$ and $f(x)$ (= either 0 or 1). Thus the quantum solution is exponentially more efficient than the classical solution.

Figure 3 shows the quantum computer solving the N bit Deutsch problem.[7] Alice prepares a physical system x in a superposition state which simultaneously occupies all possible values $[0, 2L-1]$ and another physical system y in a ground state and sends both of them to Bob. Bob performs a unitary evolution for the two input states x and y. After this evolution, y is prepared in a superposition state which simultaneously occupies all possible $f(x)$. Note that, in this way, complete quantum correlation is established between x and y. These entangled states are then sent back to Alice. Alice now imposes a π-phaseshift on y if $f(x) = 1$ and no phaseshift if $f(x) = 0$ and sends them back to Bob. This time, Bob performs the inverse unitary evolution $\hat{U}^+$ for the two inputs. After this inverse unitary evolution, y is left in the original ground state but x is either 0 or π-phaseshifted compared to the orignal state, depending on the value of $f(x)$. By appropriate decorrelation, the output x tells Alice whether $f(x)$ is all the same value (Type I) or half 0's and half 1's (Type II). This remarkable result is the consequence of the nonlocal entanglement between the two systems x and y.

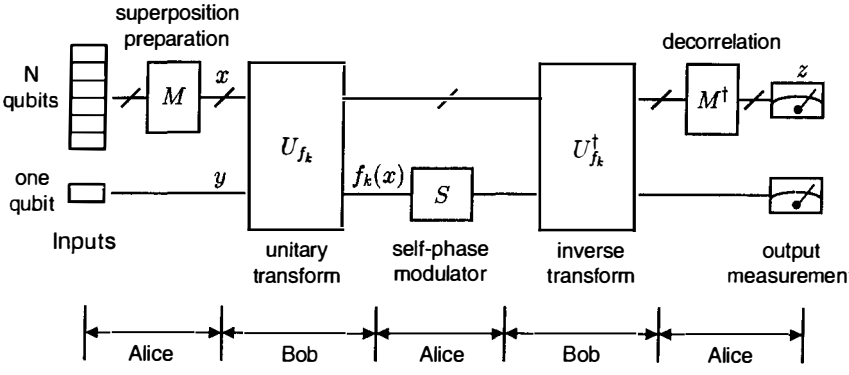


Fig. 3. A quantum computer solving the N-bit Deutsch problem.[7]

The quantum computer can be physically constructed by single photons, beamsplitters, a nonlinear Kerr medium, phaseshifters, and classical switches.[7]. As mentioned above, this example demonstrates how a certain problem can be solved efficiently by the nonlocal entanglement of superposition states.

Impossibility of the Measurement of a Single Superposition State

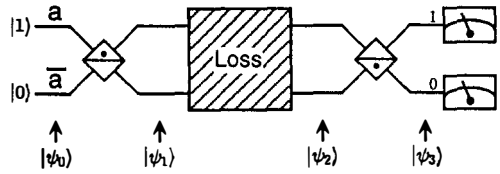
A superposition state will lose its purity once it couples to an external world. Any real physical system cannot be completely isolated from the rest of the universe, causing a serious problem for applications of a superposition state. In classical communications and computers, a binary signal (either on-off intensity or 0-π phase modulation) is distorted by attenuation (amplitude damping), dephasing (phase damping), additive noise, and imperfections of the system, which introduce a finite bit error rate. Before the bit error rate exceeds an unacceptable level, the distorted signal is detected and compared with the threshold value. A decision is then made on which signal was sent. Based on this information, a clean pulse for the next stage of communication or computation can be generated. This important concept is termed "regeneration." By accepting a finite (allowable) bit error, the signal-to-noise ratio is improved and, in this way, the bit error rate increases only additively rather than multiplicatively. In communication systems regenerators are periodically installed; in computer systems each transistor gate has this nonlinear function as its inherent property and, thus, installation of regenerators is not necessary.

Unfortunately, this concept cannot be directly applied to a quantum system because a single superposition state "cannot be measured by any means"; rather, after one particular measurement result probabilistically appears, a superposition state is destroyed (or, at least, distorted). The impossibility of measuring a single superposition state is mathematically rigorously proven and shown to be equivalent to the reduction postulate of the wavefunction.[8] As specific examples, repeated weak quantum nondemolition measurements[12] and quantum protective measurements[13] are shown to obey this rule.[8]

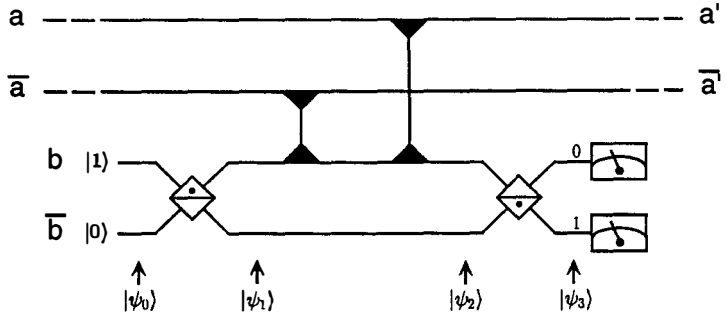
Quantum Bit Regeneration

Consider the simple optical circuit shown in Fig. 4(a). The two inputs (signal a and its conjugage $\bar{a}$) into this circuit are a single-photon state $|1\rangle$ and a vacuum state $|0\rangle$. The first beamsplitter translates the input state $|\psi_0\rangle = |a\bar{a}\rangle = |10\rangle$ into a superposition state $|\psi_1\rangle = C_0|01\rangle + C_1|10\rangle$, where $|C_0|^2 + |C_1|^2 = 1$. The single photon propagates in the upper arm with probability $|C_0|^2$ and in the lower arm with probability $|C_1|^2$. This is a dual-rail

representation of a single qubit. If both arms couple to external worlds with identical coupling strength [i.e., the loss coefficients are the same (γ) for both arms], the attenuated state $|\psi_2\rangle$ is either in the original superposition state $|\psi_1\rangle$ with probability $e^{-\gamma}$ or in the vacuum state $|0\rangle$ with probability $1 - e^{-\gamma}$. Even though the attenuation translates the originally pure signal into a mixed state, it can be decomposed into the original (intact) superposition state when a photon is not lost and into a vacuum state when a photon is lost. Therefore, the interferometer shown in Fig. 4(a) features 100% visibility if both arms have identical loss; a fact that, even classically, is well-known.



(a)



(b)

Fig. 4. (a) An optical interferometer with identical loss in two arms. (b) A quantum regeneration circuit for dual-rail qubit representation.[5)]

The quantum circuit shown in Fig. 4(b) detects whether a single photon is lost or not without destroying a superposition state.[5)] The circuit consists of two 50-50% beamsplitters and two π-phaseshifters. If there is one photon in either one of the two arms ($a\bar{a}$), the

output of the interferometer is switched. If there is no photon in both arms, the output of the interferometer is identical to the input. The presence or absence of a photon can be detected by the interferometer output. This detection process is the balanced quantum nondemolition measurement for the total photon number $\hat{a}^\dagger\hat{a}+\hat{\tilde{a}}^\dagger\hat{\tilde{a}}$ for which the superposition state $C_0|01\rangle + C_1|10\rangle$ is an eigenstate and thus is not distorted by the measurement back action.

In this way, the error caused by loss (amplitude damping) can be detected and, if the error indeed occurs, the sender can be asked to retransmit the same superposition state.

References

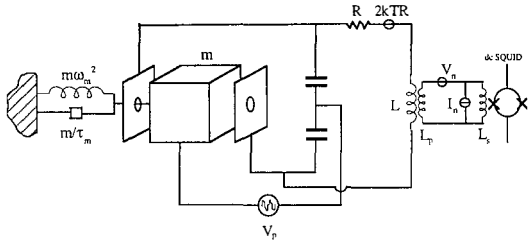
1. C. H. Bennett and D. P. DiVincenzo, *Nature* **377**, 389 (1982).
2. P. Shor, in *Proc. 35th Annual Symposium on Foundations of Computer Science* (IEEE Press, 1994).
3. P. Shor, *Phys. Rev. A* **52**, 2493 (1995).
4. A. Steane, *Proc. Roy. Soc. London A* (submitted for publication).
5. I. L. Chuang and Y. Yamamoto, *Phys. Rev. Lett.* (to be published May 13, 1996).
6. J. M. Jacobson, G. Björk, I. Chuang, and Y. Yamamoto, *Phys. Rev. Lett.* **74**, 4835 (1995).
7. I. L. Chuang and Y. Yamamoto, *Phys. Rev. A* **52**, 3489 (1995).
8. O. Alter and Y. Yamamoto, *Phys. Rev. Lett.* **74**, 4106 (1995); O. Alter and Y. Yamamoto, *Phys. Rev. A, Rapid Commun.* (to be published).
9. L. Davidovich, A. Maali, M. Brune, J. M. Raimond, and S. Haroche, *Phys. Rev. Lett.* **71**, 2360 (1993).
10. Y. Yamamoto, M. Kitagawa, and K. Igeta, in *Proc. 3rd Asia-Pacific Physics Conference*, C. N. Yang, *et al.*, eds. (World Scientific, 1988), p. 779.
11. D. Deutsch and R. Jozda, *Proc. Roy. Soc. London A* **439**, 553 (1992).
12. N. Imoto, H. A. Haus, and Y. Yamamoto, *Phys. Rev. A* **32**, 2287 (1985).
13. Y. Aharonov, J. Anandan, and L. Vaidman, *Phys. Rev. A* **47**, 4616 (1993).

A MACROSCOPIC MECHANICAL OSCILLATOR AT THE QUANTUM LIMIT LEVEL: A PRACTICAL SCHEME

E.Majorana, P. Puppo, R. Rapagnani and F. Ricci

Dipartimento di Fisica, Universita' di Roma "La Sapienza", Rome, Italy
Istituto Nazionale di Fisica Nucleare, Sezione di Roma, Rome, Italy
Dipartimento di Fisica, Universita' di Firenze, Arcetri, Italy

Back Action Evading (BAE) technique can be regarded as the classical limit of the Quantum Non Demolition measurament. We applied succefully this technique[1] in detecting the motion of a mechanical oscillator by coupling dynamically one of the quadrature components of its displacement to one of the quadrature components of the charge at the output of a readout circuit at higher resonance frequency. In this way we prevent it from the back action of the monitoring apparatus. Here we propose to optimize the detection strategy by using a dcSquid and we show that, following a classical approach, the standard quantum limit could be evaded. The detection scheme is shown in the figure.



In the BAE configuration the differential transducer is biased by an electric field $E(t) = E_s + E_n$ with

$$E_s = E_o cos\omega_e t cos\omega_m t$$

$$E_n = \frac{E_o}{2}[a_1 cos(\omega_e + \omega_m)t - \phi_1 sin(\omega_e + \omega_m)t + a_2 cos(\omega_e - \omega_m)t - \phi_1 sin(\omega_e - \omega_m)t] f \sin\omega_m t$$

where E_n include the amplitude and phase fluctuations of the field. a_1, a_2, ϕ_1, ϕ_2 are the amplitude and phase stochastic variables which have a constant power spectrum and are not cross-related. The contribution of the pump noise at the output of the circuit is reduced by the balance factor of the bridge η_z. The current flowing in the inductor L is a function of the vibration amplitude of the mechanical oscillator. To obtain performances of the system close to the quantum limit, we plan to send the signal to a dcSquid through the superconducting loop of two inductors L_p and L_s. Thus the magnetic flux at the Squid input is related to the current flowing into the inductor

L at the input of the transformer as

$$\Phi_{Squid} = M_{Squid} I_s = \frac{M M_{Squid}}{L_p + L_s} \dot{q} = T_r M_{Squid} \dot{q}$$

where I_s is the current flowing through the input inductance L_s of the Squid, $M = k\sqrt{L L_p}$ and $M_{Squid} = k_{Squid}\sqrt{L_s L_{Squid}}$ are the mutual inductances between the coil L and L_p and between L_s and the Squid inductance L_{Squid} respectively.

The power spectrum of the voltage noise V_n' that appears in the equivalent circuit, shown in the figure, is a function of the Squid voltage noise and of the Johnson noise of the bridge losses:

$$S_{V_n'} = |T_r|^2 S_{V_n} + 2k_B T R$$

The equivalent inductance L' is :

$$L' = L - T_r M$$

We assume to be able to filter out at the output all the overtone terms of ω_e and ω_m, we describe the behavior of the system in terms of the time evolution equations of the quadrature coordinates of the position of the oscillator (X_1, X_2) and of the quadrature components of the magnetic flux at the Squid input (Φ_1, Φ_2). We have:

$$\left[\frac{d}{dt} + \frac{1}{2\tau_m}\right] X_1 = -\frac{F_s}{m\omega_m} \sin\omega_m t + M_1$$

$$\left[\frac{d}{dt} + \frac{1}{2\tau_m}\right] X_2 - \frac{E_o}{4mL'\omega_e\omega_m}\Phi_2 = -\frac{F_s}{m\omega_m}\cos\omega_m t + M_2$$

$$\left[\frac{d}{dt} + \frac{1}{2\tau_e}\right] \Phi_1 + \frac{E_o}{4} X_1 = -\eta_z E_o cos\omega_m t + M_3$$

$$\left[\frac{d}{dt} + \frac{1}{2\tau_e}\right] \Phi_2 = M_4$$

We notice that these equations are coupled because of the presence of the noise terms M_1, M_2, M_3 and M_4 which include the thermal noise, the pump noise and the back action of the amplifier (the corresponding explicit formulae are reported in the reference[2]).

In the ideal case of a system of negligible pump noise, it is easy to verify that the component X_1 determines the evolution of the Φ_1 variable of the readout circuit, while the electric noise V_n' spoils both components of the electric readout, but it is squeezed only onto the homologous mechanical variable X_2. This is the typical behaviour of a BAE system.

The sensitivity of the device is derived in the case of an impulsive force acting on the mechanical oscillator. We express it in terms of the corresponding effective number of phonons (at the mechanical frequency ν_m) which are involved in the interaction process between the force and the device at the signal to noise ratio level equal one[1]. We notice that our detector is a displacement meter and that the Standard Quantum Limit is a constraint imposed to the macroscopic mechanical oscillator by the coupling with the apparatus when an amplitude-and-phase measurement is operated, then it arises from the measurement process and not from the oscillator. In this scheme the status of a quantum oscillator can carry information about classical forces that have exchanged

with it even less than one fundamental quantum of energy. For this reason we introduce the effective parameter ΔN_{eff} to express the overall sensitivity of the system by normalizing the filtered effective energy exchanged between the system and the external world to the energy quantum of the mechanical oscillator. It can be shown[1),2)] that the explicit expression of ΔN_{eff} is:

$$\Delta N_{eff} = T_n \frac{\omega_m}{\omega_e} \frac{1}{r} \left(\frac{k_B}{h\nu_m}\right)$$

where $T_n = \sqrt{S_{V_n} S_{I_n}}/k_B$ is the noise temperature of the current amplifier, k_B and h the Boltzmann and Planck constants, and r is

$$r = \frac{E_o^2}{16\omega_m \omega_e \pi m L'^2} J$$

The factor J is defined as

$$J = \int_{-\infty}^{+\infty} \frac{1}{H_o\omega^4 + H_1\omega^2 + H_2 + H_3} d\omega = \frac{\pi}{\sqrt{(H_2 + H_3)(H_1 + 2\sqrt{H_o(H_2 + H_3)})}}$$

H_o depends on the noise impedance of the amplifier $Z_n = \sqrt{S_{V_n}/S_{I_n}}$

$$H_o = \frac{1}{T_r^2 Z_n}$$

and

$$H_1 = \frac{1}{T_r^2 Z_n}\left(\frac{1}{4\tau_m^2} + \frac{1}{4\tau_e^2}\right) + \frac{1}{L'^2}\left[\frac{T_r^2 Z_n}{2} + \frac{1}{\omega_e^2 \tau_e C}\frac{T}{T_n} + \left(\frac{\eta_z E_o D}{2}\right)^2 \frac{1}{k_B T_n}(S_a + S_\phi)\right]$$

$$H_2 = \frac{1}{T_r^2 Z_n (4\tau_m \tau_e)^2} + \frac{1}{L'^2}\left[\frac{T_r^2 Z_n}{2} + \frac{1}{\omega_e^2 \tau_e C}\frac{T}{T_n} + \left(\frac{\eta_z E_o D}{2}\right)^2 \frac{1}{k_B T_n}(S_a + S_\phi)\right]$$

$$H_3 = \left(\frac{E_o}{4}\right)^2 \frac{1}{k_b T_n L'^2}\left\{\frac{k_B T}{m\omega_m^2 \tau_m} + \left[\frac{\eta_z \beta \omega_e D}{8}\right]^2 \left[S_a + \frac{S_\phi}{(2\omega_m \tau_e)^2}\right]\right\}$$

The equivalent bandwidth W of the detector can also be computed by

$$W = \frac{\int_{-\infty}^{\infty} I_{1s}(\omega) d\omega}{I_{1s}(\omega = \omega_e)} = J H_2$$

The BAE sensitivity is a function of so many parameters that for the optimization we proceede by fixing their values on the base of the present realistic technological limits. In particular, a further decrease of the e.m. pump noise contribution can be foreseen by improving the balance factor η and by lowering directly the pumping noise level. In the next we present predictions performed assuming $\eta = 10^{-7}$ (an improvement by a factor 8 in balance to the respect of what has been obtained in a previous experiment[1)]) and, concerning the pump noise level, the realistic value of -180 dBc which corresponds to the best commercially available quartz resonators[3)]. We assume to maintain the mechanical frequency as close as possible to 1 kHz and the electric one lower than 200 kHz. These two main constraints are essentially due to the limited operating conditions of

Squid-devised amplifiers . Being the system dominated by the thermodynamic component of the thermal bath, an extremely low Brownian noise level is the main requirement which can be worked out from the model. A thermodynamic temperature of few mK can be achieved by means of a $^3He - ^4He$ dilution refrigerator which has already been implemented to cool 2500 kg resonant gravitational wave antennas[4].

Regarding the mechanical oscillator, niobium resonators have been variously tested in the context of resonant gravitational wave antennas[5] and have shown very high mechanical quality factors. Niobium is superconductive at $T = 8\ K$ and is a good candidate to replace the Aluminuim alloy 5056 used in the old BAE prototype[1]. The results of the sensitivity optimization procedure are summarized in the table.

thermal bath	T [K]	Thermodynamic temperature	$1.0 \cdot 10^{-2}$
	S_{V_n} [$V^2 \cdot Hz^{-1}$]	Amplifier voltage noise	$2.0 \cdot 10^{-35}$
	S_{I_n} [$A^2 \cdot Hz^{-1}$]	Amplifier current noise	$9.4 \cdot 10^{-24}$
	S_a [Hz^{-1}]	Pump amplitude noise	$2.0 \cdot 10^{-20}$
	S_ϕ [Hz^{-1}]	Pump electric field	$2.0 \cdot 10^{-18}$
equiv. masses	m [kg]	Equivalent mass of resonator	$1.5 \cdot 10^{-1}$
	L [H]	Primary Inductor	$7.0 \cdot 10^{-4}$
Freq.	ν_m [Hz]	Mechanical frequency	1045
	ν_e [Hz]	Electric frequency	$133 \cdot 10^3$
quality factors	$Q_m = 2\pi\nu_m\tau_m$	Mechanical quality factor	$7 \cdot 10^7$
	$Q_e = 2\pi\nu_e\tau_e$	Electric quality factor	$4 \cdot 10^4$
geom. constr.	s [m]	Thickness of resonator	$3.0 \cdot 10^{-3}$
	D [m]	Gap of differerential transducer	$2.5 \cdot 10^{-5}$
coupling	E_0 [V/m]	Pumping electric field	$9.0 \cdot 10^6$
SQL	$<\Delta x>$ [m]	Standard Quantum Limit	$2.3 \cdot 10^{-19}$
detection	t_{opt} [s]	Optimum integration time	$3.2 \cdot 10^{-3}$
	ΔN_{eff}	Quantum sensitivity	0.5

We conclude that we can monitor the displacement quantity X_1 at a level corresponding to an effective energy resolution smaller than the Standard Quantum Limit.

References

1) C. Cinquegrana et al. *Phys. Rev.* **D 48**, 2478 (1993)

2) C. Cinquegrana et al. *Int. Rep., Phys. Depart.of Rome Univ.,Italy* **n.1034** (1994)

3) Wenzel Associated, Ltd.

4) P. Astone et al. *Europhysics Letters* **16**, 231 (1991)

5) D.G. Blair et al. *Phys. Rev. Lett* **74**, 1908 (1995)

UN OSCILLATEUR MÉCANIQUE MACROSCOPIQUE AU NIVEAU DE LA LIMITE QUANTIQUE: UN SCHÉMA OPERATIF.

Résumé

Avec un system à Évasion de la RétroAction (ERA) il est possible de contourné la limite de sensibilité des détecteurs linéaires de déplacement. Nous présentons ici les prévisions sur les performances d'un système ERA optimisé par rapport au prototype que nous avons déjà mis en fonction en régime classique. Nous montrerons que ce système expérimental peut arriver au delà de la Limite Quantique Standard.

NEW ASPECTS OF NEUTRON OPTICS

Robert Golub°, Roland Gähler+, Joseph Felber+, Steve Lamoreaux*

°Hahn Meitner Institute

14109 Berlin, Fed. Rep. of Germany

+Technical University München

85748 Garching, Fed. Rep. of Germany

*University of Washington, Seattle, WA 98 195, USA

Neutron optics with time dependent devices, *e.g.* moving slits and mirrors, offers new possibilities of investigating, among other things, details of the time dependent Schrödinger equation for massive particles, the transition from quantum to classical behavior for a simple system and the coherence properties of neutron beams.

Spin optics is a kind of analogue computer for optical effects where the average spin of *classical* particles arriving at a point can serve as a model for the optical excitation in ordinary optics displaying interference and diffraction effects. We give an introduction to the concept and mention some possible applications.

1 Introduction

Neutron optics with time dependent devices, *e.g.* moving slits and mirrors, offers new possibilities of investigating, among other things, details of the time dependent Schrödinger equation for massive particles, the transition from quantum to classical behavior for a simple system and the coherence properties of neutron beams. We present the results of calculations of the passage of a neutron beam through a rapidly moving slit including the "two slit interference in time" where the components of the wave function passing through the slit at one moment in time interfere with those passing through at a later opening time.

We then show experimental results for neutron reflection from a vibrating mirror where quantized energy transfer is observed and the tendency to classical behavior as the vibration amplitude is increased is demonstrated.

We go on to discuss some questions associated with coherence properties of neutron waves and show that a sideband interferometer (based on interference between the different waves reflected from a vibrating surface) can be used to study and/or produce both longitudinal and transverse coherent effects.

Several of the effects discussed here have been observed, or are expected to be observed shortly, with atom beam optics (see the contribution of Andrew Steane to these proceedings).

Spin optics is a kind of analogue computer for optical effects where the average spin of *classical* particles arriving at a point can serve as a model for the optical excitation in ordinary optics displaying interference and diffraction effects. We give an introduction to the concept and mention some possible applications. The idea is that in the presence of a uniform magnetic field the Larmor precession angle for different particle trajectories is a measure of the path length followed by the different particles and the average spin of all the particles arriving at a given point is a vector superposition of the individual spins of these particles. This superposition of contributions with different phase shifts gives rise to interference and diffraction effects with macroscopic wavelengths.

2 Time dependent neutron optics

2.1 Fast chopping of a neutron beam - the 'quantum chopper'

If a neutron or other massive particle beam goes through a fast chopper, *i.e.* a chopper whose opening time, T, is not too much longer than $1/\omega_o$, where $\hbar\omega_o$ is the particle kinetic energy, then the resulting beam will have its energy spectrum broadened by $\Delta\omega \sim \hbar/T$ and this can be observed by usual spectrometric methods.

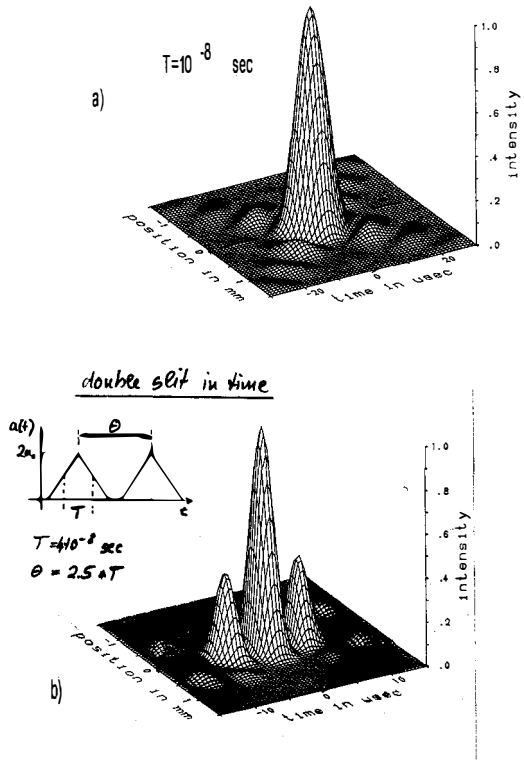


Figure 1: a) Detector response as a function of position and time for a triangular slit opening function with T=10^{-8} sec. b) Detector response for a double pulse triangular slit function.

On the other hand if the detector intensity is measured as a function of time there will be interesting interference effects between the different energy components of the beam. These effects are analogous to the diffraction and interference effects seen in ordinary optics and have been referred to as diffraction in time.[1)].

In addition to the diffraction in time we have considered a simultaneous spatial diffraction from the time dependent slit.[2)]. The calculation is based on the Green's function for the time dependent Schrödinger equation. The intensity arriving at a detector positioned at a distance L from the slit is given by

$$\begin{aligned} \psi(x=d, y=0, z=L, t) &= \psi_+ - \psi_- \\ \psi_\pm &= -\frac{\hbar k}{2\pi m}\frac{e^{-i\omega_o t}}{d}\int_{0_-}^{t} d\tau \exp i\left[\omega_o \tau + K_\pm(\tau)/\tau\right] \\ K_\pm(\tau) &= \omega_o \tau_o^2\left[1+\frac{d^2}{2} \pm 2da\,(t-\tau)\right] \end{aligned} \tag{1}$$

where the beam is travelling in the $+z$ direction, the positions of the slit boundaries at time t_o are given by $(x_o = \pm a(t_o), y, z_o = 0)$, $\tau_o = L/v$, and $\tau = t - t_o$, is the travel time between slit and detector. In fig. 1) we show the expected detector intensity as a function of time and detector position (distance perpendicular to the beam axis $=d$) for an experimentally realizable triangular slit opening function, a neutron wavelength of $25\mathring{A}$, $L = 1.5m$ and a maximum slit opening of $10\mu m$.

Fig 1a) shows the fully developed interference pattern in space and time, where T, the time for the slit to fully open was taken as 10^{-8} sec. Fig 1b) shows the pattern to be expected for the case of two successive triangular slit openings each with $T = 10^{-8}$ sec, separated in time by 2.5×10^{-8} sec. The similarity with the usual two slit interference pattern is clear.

Experiments with the goal of observing these effects are in progress. Opening times of $T \approx 3 \times 10^{-8}$ sec have been achieved.

2.2 Reflection from a Vibrating Surface - the transition from quantum to classical behavior

There have been several suggestions that quantum mechanics might have to be altered somewhere between the region where physical behavior is 'pure quantum'-like and the region where physical behavior is governed by purely classical physics.[3)]. We present a system where the transition between quantum and classical behavior is rather straightforward and where the observational possibilities are good. However our work does not address the issue of wave function collapse at the detector. We have studied the transmission and reflection of slow neutrons at a mirror whose position undergoes a high-frequency oscillation in the direction perpendicular to its surface. The mirror is of macroscopic dimensions and can be regarded as a classical object,

analogous to the oscillating magnetic field in magnetic resonance experiments. We show both theoretically and experimentally that the interaction with the neutrons is quantized and can be interpreted as a coherent multiphonon exchange. The number of exchanged phonons increases with the oscillation amplitude and the behavior gradually approaches that of a classical particle bouncing from a moving surface.

The reflection of a particle from a constant 1-dimensional potential step is one of the first problems we learn to solve in quantum mechanics. For a one dimensional potential step of height V located at $x = a$, the reflection coefficient is given by $R = R_o e^{2ika}$ where R_o is the reflection coefficient for a potential step located at $x = 0$ ($R_o = 1$ for the case of total reflection, $E < V$, where E is the energy of the incoming particle). For a vibrating wall

$$a = \epsilon \sin(\omega_v t + \varphi) \tag{2}$$

where ω_v is the vibration frequency, and for $\omega_v \ll E/\hbar$ we can write

$$R(t) = R_o e^{2ikt\sin(\omega_v t+\varphi)} = R_o \int G(\omega) e^{i\omega t} d\omega \tag{3}$$

with

$$G(\omega) = \sum_n e^{in\varphi} J_n(2k\epsilon)\, \delta(\omega - n\omega_v) \tag{4}$$

which is exact for $V \to \infty$. For finite V it is necessary to resort to numerical solutions, but the basic behavior is similar to that shown in equ. [4]. What we have is a phase modulated neutron wave with a spectrum familiar from frequency modulation. We see that the amplitude for an energy transfer $n\hbar\omega_v$ is given by $J_n(2k\epsilon)$, the n^{th} order Bessel function of the first kind.

For a classical particle colliding with a moving wall the particle velocity after the collision is given by

$$v_2 = v_1 + 2v_w(t), \quad v_w(t) = \omega_v \epsilon \cos(\omega_v t + \varphi) \tag{5}$$

with v_1, the initial and v_2 the final velocities of the particle. v_w is the wall velocity.

Then the probability distribution for v_2 is given by

$$P(v_2)\, dv_2 \propto dt = \frac{dv_w}{dv_w/dt} = \frac{dv_w}{\omega_v^2 \epsilon \sin(\omega_v t + \varphi)} \propto \frac{dv_2}{\sin\left[\cos^{-1}(u/\omega_v\epsilon)\right]} \tag{6}$$

with $u = (v_2 - v_1)/2$. This distribution is, of course continuous, and reaches peaks at $u = \pm\omega_v\epsilon$. It is rather flat in the region of $u = 0$. According to [5] u is limited to values $|u| < \omega_v\epsilon$. This distribution is only valid in certain cases. For some choices of parameters there exist trajectories which make multiple collisions with the wall and which approach a chaotic behavior.[4)]

In fig. 2) we show a comparison of the quantum mechanical and classical energy spectra for totally reflected particles for different values of the modulation index, $\alpha = 2k\epsilon$ and the relative wall velocity $\gamma = \omega_v\epsilon/v_1$. The approach to the classical distribution with increasing α is evident. It is interesting to note that the classical limit is not approached by all spectral components in a uniform way but is only reached by averaging over many spectral lines when the energy resolution of the measurement is much wider than the separation between lines: $\hbar\omega_v$.

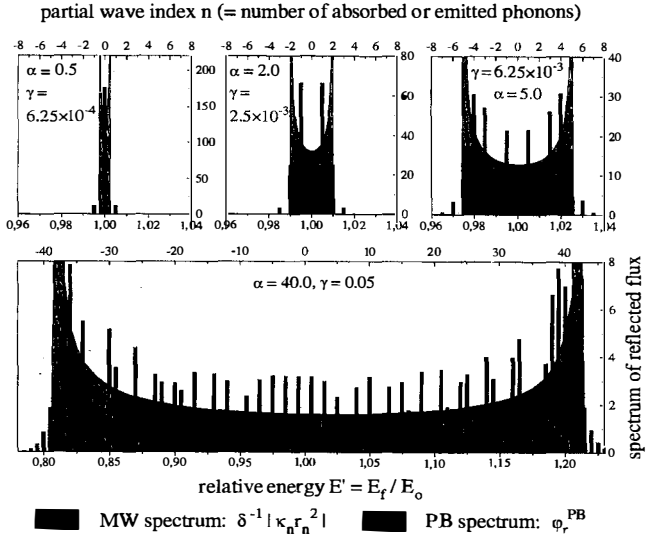


Figure 2: Quantum (MW, matter wave) and classical (PB, particle beam) energy spectra (calculated) of particles reflected from a vibrating surface. Parameters are defined in the text.

This is even clearer in fig. 3) where we show the results of calculations for a larger modulation index. In fig. 4) we show the results of a series of measurements made at a neutron beam at the research reactor in Geestacht (FRG) Germany. The incident neutron velocity was 164 m/sec ($E_o = 140\mu ev$) and the velocity component perpendicular to the mirror surface was 5.7 m/sec ($E_\perp = 171nev$). The potential height of the nickel coated glass mirror was $V = 235nev$. The grey bars show the calculated positions and intensities of the various sidebands corrected for the resolution of the apparatus and inhomogeneities in the mirror oscillation which were measured by optical measurements of the surface modes using a Michelson type interferometer. The line separation was about $3 \times 10^{-4} rad$ in the case of $\omega_v = 2\pi \cdot 693kHz$, corresponding to a distance of $0.8mm$ at the detector plane. The experiment was performed using a two-dimensional position sensitive detector which was developed at Garching[5)] and which had a resolution of $35 \times 35\mu m^2$, an efficiency of about 90% at 24Å and very low background and γ sensitivity.

While the trend to a classical-like behavior is clear we were not able to go into the highly classical regime in this first experiment. We also see the influence of the resolution on the approach to the classical limit. The measured energy splitting of the reflected beam ($2.86nev$)agreed with that expected from the applied frequency ($693kHz$) within the error of $10^{-2} nev$.

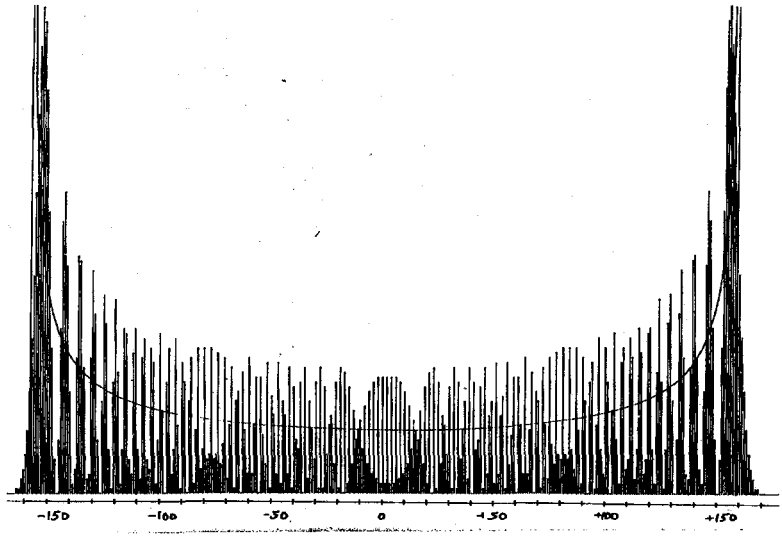


Figure 3: Same as fig. 2 but for a larger modulation index $\alpha = 73$ and $\delta = \omega_v/\omega_o = 10^{-3}$ ($\hbar\omega_o$ = particle kinetic energy). The complicated nature of the approach to the classical spectrum (smooth curve) is clear.

3 Coherence properties of neutron beams and a side-band interferometer

The question of the coherence properties of neutron beams arises in connection with the neutron interferometer and has been the subject of some discussion in the literature[6]. A neutron interferometer measures

$$I(\Delta x) = \int dx \psi(x) \psi(x + \Delta x) = \int dk \, |A(k)|^2 e^{ik\Delta x} \tag{7}$$

where $A(k)$, the Fourier transform of $\psi(x)$. is the amplitude of the momentum k in the beam, so that the so-called 'coherence length', Δx, contains no more information than the spectral width, *i.e.* it is the same for a mixture as for a superposition of different $k's$. Now there is a real physical difference between a superposition and a mixture so the question arises as to how we can make the distinction in this case.

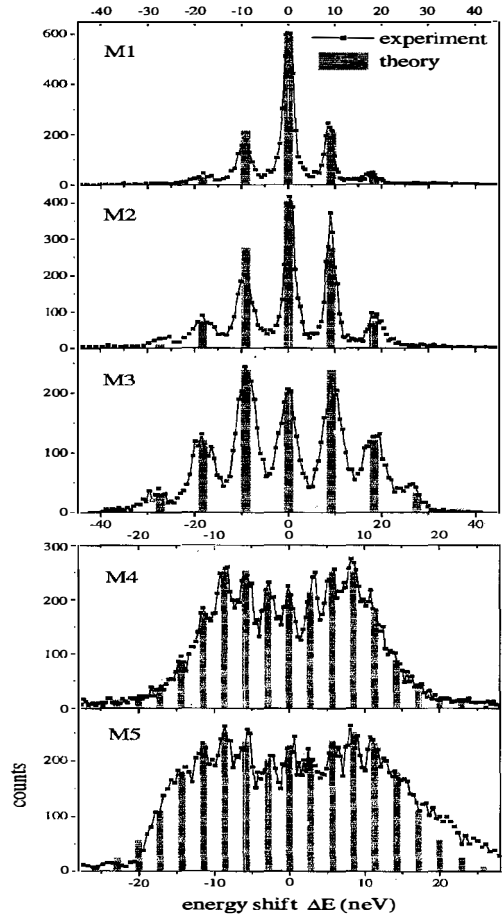


Figure 4: Comparison of measured and calculated energy spectra for neutrons reflected from a vibrating surface: M1-3: vibration frequency 2.2206 MHz. Mean vibration amplitudes: M1 5,3 nm ($\alpha = .95$), M2 8.0 nm ($\alpha = 1.45$), M3 10.9 nm ($\alpha = 10.9$). M4, M5 frequency 692.95 kHz, amplitudes 23.2 and 32.1 nm ($\alpha = 4.21$ and $\alpha = 5.82$).

To do this we must measure the average phase difference $\langle \varphi_{ij} \rangle$ between the waves having momenta $\overrightarrow{k_i}$, $\overrightarrow{k_j}$, or the average value of the auto-correlation function in $\overrightarrow{k}$ space: $\left\langle A^* \left(\overrightarrow{k_i}\right) A \left(\overrightarrow{k_j}\right) + c.c. \right\rangle$,

i.e. we must form an interference or a beat between different $\vec{k}$. In optics, beams with $\Gamma\left(\vec{k_j}-\vec{k_i}\right) \equiv \left\langle A^*\left(\vec{k_i}\right) A\left(\vec{k_j}\right)\right\rangle = \delta\left(\vec{k_j}-\vec{k_i}\right)$ are called 'chaotic'. As seen from equ.[8] this implies that the ensemble average $\left\langle\left|\psi_{inc}\left(\vec{x}\right)\right|^2\right\rangle$ is independent of x.

It is usually assumed that neutron beams share this property[7], and although the authors see no reason to question this it still should be addressed as an experimental question. We now give a description of one way in which this can be done. In the event that neutron beams should prove to be 'chaotic' our technique will allow the production of beams with $\Gamma\left(\vec{q}\right) \neq \delta\left(\vec{q}\right)$ which can then be further investigated in what we call the 'side-band interferometer.

For example, a periodic system with reciprocal lattice vector $\vec{q}$ can scatter an incoming momentum $\vec{k_i}$ to a final momentum $\vec{k_i}+\vec{q}=\vec{k_2}$ and a momentum $\vec{k_j} \rightarrow \vec{k_j}-\vec{q}=\vec{k_3}$. Now for $\vec{k_j}-\vec{k_i}=\Delta\vec{k}=2\vec{q}$ we have $\vec{k_2}=\vec{k_3}$ and the scattered intensity will contain a term proportional to[8]

$$\int d^3k \left\langle A^*\left(\vec{k}\right) A\left(\vec{k}+2\vec{q}\right)\right\rangle = \int d^3x \left\langle\left|\psi_{inc}\left(\vec{x}\right)\right|^2\right\rangle e^{-i2\vec{q}\cdot\vec{x}} \tag{8}$$

Now for scattering from static systems the energy, and thus $\left|\vec{k}\right|$ is unchanged so that if we wish to investigate the coherence between $\vec{k_i}$, $\vec{k_j}$ which are parallel, *i.e.* longitudinal coherence, we have to scatter from a time dependent system in order to produce changes in the energy of the scattered neutrons. As shown above reflection from a vibrating surface can produce coherent beams of different energies. Thus reflection of a beam from a surface excited by two frequencies or, what is more practical, successive reflection from two surfaces vibrating at frequencies ω_v, $\omega_v'=\omega_v+\Omega$ with $q\left(q'\right)=\frac{m}{\hbar k}\omega_v\left(\omega_v'\right)$ will take say $\vec{k_i} \rightarrow \vec{k_i}+\vec{q}$ and $\vec{k_j} \rightarrow \vec{k_j}+\vec{q'}$. The final momenta will be equal in the case where $\vec{k_j}-\vec{k_i}=\vec{q}-\vec{q'} \equiv \vec{\delta}=\frac{m}{\hbar k}\Omega$. In this case the scattered intensity will contain a term proportional to $\left\langle\int d^3k A^*\left(\vec{k}\right) A\left(\vec{k}+\vec{\delta}\right)\right\rangle$ where δ is proportional $\Omega=\omega_v'-\omega_v$ and, since we have an energy change due to the reflection, $\vec{\delta}$ can be parallel to $\vec{k}$ so that we can investigate longitudinal as well as transverse coherence. In a similar way, it is easy to see that we can produce interferences between different side bands in the case where we have a monochromatic incident beam (side-band interferometer). The point is that the different reflected beams shown in fig. 4) above are all coherent.

If we wish to distinguish between a state and a mixture we have to find some way of measuring the phase between different components of the incident state, *i.e.* we need a method of getting these components to interfere with each other. In the work under discussion, this is brought about because the individual components, with slightly different incident energies can scatter inelastically into the same final energy state where they can beat to form a steady contribution to the detected intensity. It is easy to see that the same phenomenon can take place if the incident beam is passed through a fast chopper as in sec. **2.1** above. In this case the spectra resulting from each incident $\vec{k}$ will overlap and interfere with each other if there is a correlation between the different $A(k)$. We hope to discuss this in more detail at a later date.

Experiments of this type are also possible using atomic beams and some very beautiful experiments have been described by Andrew Steane at this conference.

4 Neutron Spin Optics

Using Larmor precession of particle (neutron) spins and measuring the average spin of particles arriving at a point by various paths it is possible to recreate many optical phenomena with effective wave lengths which can be macroscopic in size. In classical optics, interference and diffraction effects are often described on the basis of the Huygens-Kirchoff principle - summing up the optical vibrations emanating from different source points of a diffracting structure such as a collection of holes or slits. The key features are

1. that the phase of the contribution from each source point at a given observation point is a linear function of the distance between the two points.
2. the signal amplitude is a linear superposition of the individual contributions.
3. the detected intensity is proportional to the square of the resulting amplitude.

Larmor precession in a constant magnetic field B shares properties 1) and 2)[9]. We consider the case where the spins are precessing in a plane perpendicular to the field, in which case they remain in this plane. The precession frequency $\omega_L = \gamma B$ and the distance traveled in one period is $\lambda_L = v2\pi/\omega_L$. $\lambda_L = 1mm$ for a particle velocity, $v = 10^3 m/$ sec and $B = 300$ *gauss*. Note that in the cases of interest λ_L is much greater than the de Broglie wavelength of the particles so that they can be considered to be moving classically. That is the particles will move in straight lines through any slit large enough to show optical effects on the scale of λ_L, unless we fill the slit with scattering centers which scatter the particles into a reasonably large angle. (Another variant is to have a source emitting polarized particles in a wide angle directly behind the slit). In either case the particles will be assumed to arrive at the slit all polarized in the same direction (x direction, while they travel in the y direction). For particles which travel a distance r_i between the slit and a detector screen the spins will precess through an angle $\varphi_i = k_L r_i$ where $k_L = 2\pi/\lambda_L$. The spin of the i^{th} particle arriving at a point P in the detector plane can be represented by $\overrightarrow{\sigma_i}(P) = e^{i\varphi_i}$ and the average spin of all the particles arriving at P will be given by

$$\langle \overrightarrow{\sigma_i}(P) \rangle \propto \sum_i e^{ik_L r_i} \tag{9}$$

In the case where the source plane has, say, two slits, the sum [9] will be the same as the amplitude of the usual two slit interference pattern corresponding to a wavelength λ_L. This is the case in spite of the fact that there is nothing, neither particles nor waves, which goes through both slits. Note that since we start with particles all having the same polarization the particles

are effectively all coherent. In order to produce the interference pattern we sum the signal [9] over a long time. In addition the whole system can be described classically.

Using this technique we can construct the spin optics analogues of diffraction gratings, Fresnel lenses, holograms and other optical devices with wavelengths in the microwave region.

If it was possible to use spin optics as a new scattering technique to study materials we would be able to carry out neutron scattering studies at macroscopic (mm to cm) distances. Unfortunately, as opposed to the usual scattering techniques where $\langle E_i^2 \rangle$ is measured (E_i refers to the amplitude of the radiation in question) with spin optics we can only measure $\langle \overrightarrow{\sigma_i}(P) \rangle^2$. This is easily seen because $\left\langle (\overrightarrow{\sigma_i})^2 \right\rangle = 1$ for any density matrix. This means that spin optics scattering from a collection of scatters randomly placed with respect to each other will tend to cancel because of the random phases associated with the positions of the scatterers. Thus if we consider a scattering system consisting of a series of grains, each containing a number of incoherent scattering centers, we will not have any signal unless the grains are arranged in a coherent manner (say on a periodic lattice) in which case we could measure grain sizes on the scale of the spin optics wavelength λ_L which is much larger than length scales available with classical small angle scattering. Another possibility is to work with a single grain which may be possible given the macrocopic size of λ_L.

At the moment the authors see no other possibilities of overcoming this problem, which can be characterized as the lack of a square-law detector for spin optics. In addition we do not know of any situation where the above mentioned technique might be useful. We look forward hopefully to suggestions from our readers.

5 REFERENCES

[1] M. Moshinsky, Phys . Rev. 88, 625 (1952).

[2] R. Gähler and R. Golub, Z. Phys. B56, 5 (1984),

J. Felber, G. Müller, R. Gähler and R. Golub, Physica B162, 191 (1990),

J. Felber, R. Gähler and R. Golub, Physica B151,135 (1988).

[3] There is a large body of literature dealing with this question. See, e.g., the articles by O. Penrose, E.J. Squires, L.Diósy, P. Pearle and G.C. Ghirardi in *Quantum Chaos-Quantum Measurement,* Vol. 158 of *NATO Advanced Study Institute, Series B: Physics,* ed. by P. Cvitanović, I. Percival and A. Wirzba (Plenum, 1992), and R. Penrose, *The Emperor's New Mind* (Oxford Univ. Press, 1989) for a more popular treatment of these ideas.

[4] J. Felber, PhD. Thesis, Physics Dept., Technische Universität München, 1994.

[5] C. Rausch *et al*, SPIE 1737, 255 (1992)

[6] G. Comsa, Phys. Rev. Lett. 51,1105 (1983),
H. Kaiser, S.A. Werner and E.A. George, Phys. Rev. Lett. 51,1106 (1983),
H.J Bernstein and F.E. Low, Phys. Rev. Lett. 59, 951 (1987).

[7] V.F. Sears, *Neutron Optics,* Oxford, (1989).

[8] R. Golub and S.K. Lamoreaux, Phys Lett. A162,122 (1992).

[9] R. Gähler and R. Golub, Phys. Lett., to be published (1996).

QUANTUM ERROR CORRECTION

Andrew M. Steane
Clarendon Laboratory, Oxford OX1 3PU, England.

The concept of *quantum information* is proving to be very useful in efforts to elucidate the nature of quantum mechanics. Both quantum communication and quantum information processing has been shown to be fundamentally different from its classical counterpart. Examples where this difference is highlighted are secure key distribution for cryptography, and the existence of fast algorithms for an idealised quantum computer. However, actual attempts to realise these possibilites run up against a further fundamental part of quantum mechanics: the problem of the instability of coherence.

All physical systems are subject to random fluctuations, including those degrees of freedom which may be described in terms of classical mechanics. However, classical degrees of freedom may be stabilised to a very high degree, either by making the ratio of system size to perurbation size very large (passive stabilisation), or by continuously monitoring the system and providing greatly enhanced 'inertia' against random fluctuation be means of feedback control (active stabilisation). Of these two possibilities, the former, that is passive stabilisation, can be applied in the quantum regime only by making the perturbations small rather than by making the system large, and stabilisation beyond a certain degree is in practice ruled out. It is not hard to show that this makes the experimental realisation of a quantum computer of useful computational power impossible by any currently attemptable method. The physics of quantum information processing remains interesting, one should add. It is simply that great computing power is not available.

The method of *active* stabilisation is extremely powerful in classical systems, and is at the heart of mechanical devices from early steam engines to the latest microchip processors. However, it is not obvious whether anything like active stabilisation is possible for a quantum system, since feedback control involves dissipation, and therefore is non-unitary. Hence one may frame the following question:

> *"Is active stabilisation of a quantum bit possible?"*

The idea of a quantum bit or qubit is introduced in order to emphasize that the aim is to stabilise a complete quantum state, not just a chosen observable. Also, we are concerned with the properties of the quantum state, not with the physical system expressing it. For example, a single qubit may be expressed by a system whose Hilbert space has many more than two dimensions. Among the possible changes such a system may undergo, some will affect the stored single qubit of quantum information, but others will not.

The surprising answer to our question is "yes," with some important *provisos* which depend on the type of stabilisation sought. The stabilisation is based on the classical theory of error correction, which provides a very powerful technique by which classical information can be transmitted without errors through the medium of a noisy channel[1]. Classical error correction operates by the judicious use of redundancy, that is, sending the same information many times. In this sense it is akin to making the system larger in order to make it more resistent to perturbations. However, the precise way in which the redundancy is introduced is very important. The type of redundancy, or *encoding*, employed must be carefully matched to the type of noise in the channel. Typically, one considers the case of random noise which affects different bits independently, but this is not the only possible case. The encoding enables the most likely errors in the information to be identified and corrected. This corrective procedure is akin to active stabilisation, and brings the associated benefits of powerful noise suppression.

To understand the application of these ideas to the quantum regime, it is best to start with a simple example. Thus, suppose we have a collection of spin-half particles, each of which is subject independently to random 'flips' or amplitude errors $|0\rangle \leftrightarrow |1\rangle$, but which otherwise is stable (in particular, the precession is free of phase error). Whenever such a flip occurs, the relevent two-state system may become entangled with its environment. In order to stabilise a single qubit, in the general state $a\,|0_L\rangle + b\,|1_L\rangle$, we express it by means of three two-state

systems, with the 'encoding' $|0_L\rangle \rightarrow |000\rangle$, $|1_L\rangle \rightarrow |111\rangle$. Thus the total initial state of the three spins is $a|000\rangle + b|111\rangle$. After a period of time, during which random flips may occur, the three-spin system is measured twice. The first measurement is a projection onto the two-state basis

$$\{|000\rangle + |111\rangle + |001\rangle + |110\rangle\,, \quad |010\rangle + |101\rangle + |100\rangle + |011\rangle\} \tag{1}$$

The second measurement is a projection onto the two-state basis

$$\{|000\rangle + |111\rangle + |010\rangle + |101\rangle\,, \quad |001\rangle + |110\rangle + |100\rangle + |011\rangle\} \tag{2}$$

Each measurement has two possible results, which we will call 0 and 1. Depending on which results R are obtained, an appropriate action is carried out: if $R = 00$, do nothing; if $R = 01$, flip the rightmost spin; if $R = 10$, flip the middle spin; if $R = 11$, flip the leftmost spin. If, during the time interval when the system was left to evolve freely, no more than one spin flipped, then this procedure will return the three-spin state to $a|000\rangle + b|111\rangle$. It is remarkable that this can be done without gaining information about the values of a and b and thus disturbing the stored quantum information. During the correction procedure, the entanglement between the system and its environment is transferred to an entanglement between the measuring apparatus and the environment. The qubit is actively isolated from its environment by means of this carefully controlled *entanglement transfer*.

The above error correction technique is based on the simplest classical error correcting code. More advanced techniques can be deduced from more advanced known classical codes, and the following striking results are obtained. First, completely general error processes can be corrected, including relaxation and entanglement with the environment, and to do this it is sufficient merely to be able to correct for spin flips (σ_x spin operator) and sign flips (σ_z spin operator)[2–5]. Second, a subset of the classical codes can be adapted directly to the quantum context[3–5]. Third, the probability of failure of the quantum error correction falls exponentially with the redundancy, in the limit of large redundancy, as long as the error rate is below a given level[4,5]. Finally, the degree of required redundancy and the complexity of the correction process rises only as a low-order polynomial function of the number of qubits to be corrected[5].

Note that we define the term 'error' to mean in general *any* contribution to the evolution of a quantum system which is unpredictable. Usually therefore the errors will be continuous rather than discrete, and will affect all the qubits rather than a subset. However, during

error correction the system is projected onto a subspace of its Hilbert space which contains only state vectors with a specific error syndrome. Therefore the continuous error process is rendered discrete (collapsed) by the projective measurement.

The main *proviso* to all the above is that the correction process can itself be carried out without errors. This is clearly a huge assumption. It is probably reasonable in the context of quantum communication[7,8], since there one eventually wishes to measure the communicated qubits, and the bulk of the error correction can be carried out on the classical information obtained *after* the qubits are measured. The context of quantum computing is another matter, however, and it remains to be seen whether quantum error correction can be made sufficiently robust against noise during the correction process itself. Thus quantum theory may still rule out the possibility of a powerful quantum computer.

The author is supported by the Royal Society.

1. MacWilliams F. J. & Sloane, N. J. A. 1977 *The Theory of Error-Correcting Codes.* Amsterdam: North Holland.
2. P. W. Shor, Phys. Rev. A **52**, R2493 (1995).
3. A. M. Steane, submitted to Phys. Rev. Lett.
4. A. R. Calderbank and P. W. Shor, submitted to Phys. Rev. A.
5. A. M. Steane, submitted to Proc. Roy. Soc. A.
6. Related preprints now in circulation are those of I. L. Chuang and R. Laflamme; R. Laflamme, C. Miquel, J. P. Paz and W. H. Zurek; C. Macchiavello and A. Ekert.
7. C. H. Bennett, G. Brassard, S. Popescu, B. Schumacher, J. A. Smolin, and W. K. Wootters, Phys. Rev. Lett. **76**, 722 (1996).
8. A. Ekert, R. Jozsa, C. Macchiavello, S. Popescu, A. Sanpera and W. K. Wootters, Oxford University preprint.

Résumé

Pour des systèmes classiques, les effets indésirables du bruit peuvent être contrôlés par une stabilisation active. Dans le cas des systèmes quantiques, pour lesquels une evolution unitaire doit être préservée, une telle stabilisation n'est pas possible. Ceci semble annéantir les espoirs que l'on fonde sur les possibilités d'un ordinateur qui fonctionnerait de manière quantique. Cependant, en adaptant des méthodes classiques de correction d'erreur, l'information quantique peut être stabilisée activement. La communication d'états quantiques en presence de bruit est ainsi rendue possible, et il est probable que l'élaboration de calculs au niveau quantique puisse bénéficier de ces nouvelles techniques.

CLOCK SYNCHRONISATION WITH QUANTUM FIELDS

Marc Thierry Jaekel[a] and Serge Reynaud[b]
(*a*) Laboratoire de Physique Théorique, ENS, CNRS, UPS,
24 rue Lhomond, F-75231 Paris Cedex 05, France
(*b*) Laboratoire Kastler Brossel, ENS, CNRS, UPMC,
4 place Jussieu, F-75252 Paris Cedex 05, France

Light signals play a central role in introductory presentations of relativity theory. As well known, simultaneity is not an absolute notion for events occuring at different spatial positions. Remote clocks can be synchronised by transfering field pulses from one clock to the other one [1]. The rates of the clocks are determined by frequency standards and can also be compared by transfering a light pulse used as a frequency reference rather than a time reference.

The concepts of time or frequency references certainly differ from the classical notions of time or frequency. The time reference cannot be identified with the time coordinate which is used to write evolution equations of classical or quantum field theory. Similarly, the frequency reference cannot be identified with the frequency parameter which is used to distinguish the various field modes. Time and frequency references are physical observables delivered by specifically designed physical systems, namely clocks and frequency standards. As well known, the modern metrological definition of time and frequency has its roots in atomic physics [2]. More generally, clocks and frequency standards belong to the quantum domain while light signals are built with quantum fields. Any practical realisation of time or frequency has therefore to meet quantum limitations [3]. The aim of the present paper is precisely to discuss the questions which arise when clock synchronisation techniques are confronted with the quantum nature of observables [4]. Problems associated with the generation of time by quantum clocks will not be adressed here.

In order to specify these questions, it is worth refering to a *gedanken* experiment, where an emitter sends a pulse containing some time or frequency information to a remote receiver. In classical physics, the problem of encoding time or frequency information in a pulse and then retrieving it has well known solutions. The reference to be shared by the remote observers is chosen as a quantity preserved by field propagation, that is either the light-cone variable u or the field frequency ω

$$t_e - x_e = u = t_r - x_r \qquad \omega_e = \omega = \omega_r \tag{1}$$

t_e and t_r are the emission and reception times, as delivered to the emitter and receiver by their private clocks; x_e and x_r are the space coordinates of the emitter and receiver, as measured along the line of sight; ω_e and ω_r are the frequency standards which control the clock rates; for simplicity, we have supposed the two clocks to be identical. Now, the discussion of the same *gedanken* experiment in quantum physics raises serious difficulties, since the formalism does not directly supply definitions of time or frequency observables. In order to make an operational discussion of clock synchronisation possible, we have first to give a quantum definition of the time reference, although it is commonly stated that there is no time operator available in the quantum formalism. Such a quantum definition is expected to show quantum fluctuations of time conjugated to quantum fluctuations of energy, in consistency with the fourth Heisenberg inequality. We have in fact to define a time and a frequency observable for each field pulse. The quality of the time and frequency transfer procedures is optimized by using a field pulse having respectively a short duration or a good spectral purity. Since these requirements cannot be satisfied simultaneously with an arbitrary precision, clock synchronisation procedures have to meet quantum limitations. Though these limitations cannot be considered to be of practical importance in present time-frequency metrology, their precise estimation appears to be important from the point of view of fundamental physics [5].

A further question now arises, concerning the consistency of time and frequency definitions with frame transformations. In classical relativistic physics, time and frequency are shifted in a frame transformation. The consistency of these shifts is ensured by Lorentz invariance for transformations from an inertial to another inertial frame while they are ensured by covariance rules for transformations from an inertial to an accelerated frame [6]. The quantum definitions of time and frequency references have also to be affected

in frame transformations. We may expect that Lorentz transformations affect in the same manner quantum and classical definitions, since quantum field theory obeys Lorentz invariance exactly as classical field theory. However, we cannot know a priori how transformations to accelerated frames affect the quantum definitions since we cannot derive directly quantum shifts from classical covariance rules. In the present paper, we will derive the quantum shifts from the underlying conformal symmetry of quantum fields, and then show that the consistency between time- and frequency-shifts follows from this symmetry.

Electromagnetism in four-dimensional spacetime is invariant not only under Lorentz transformations, but also under the larger group of conformal transformations [7]. These transformations are known to fit the relativistic definition of uniformly accelerated motion [8]. Light rays remain straight lines with such a representation of accelerated frames while frequencies are preserved during light propagation. It has also been shown that the definitions of vacuum [9] and particle number [10] are unchanged under conformal transformations, so that these concepts may be safely manipulated by inertial and accelerated observers. This is consistent with the seminal discussions of Einstein who proposed in 1905 the hypothesis of light quanta [11] and introduced, a few months later, the principle of relativity. In the latter paper [1], he noticed that energy and frequency of an electromagnetic field change in the same manner in Lorentz transformations, thus implicitly referring to the consistency of the hypothesis of light quanta with the principle of relativity. In modern quantum theory, the similarity of energy and frequency shifts has to be interpreted as an invariance property for particle number. This property is well known for Lorentz transformations, but it is usually not admitted for transformations to accelerated frames. The latter are commonly represented by Rindler transformations [12] which do not preserve the propagation equations of electromagnetic fields, so that light rays appear curved and frequencies changed during propagation. Rindler representation of accelerated frames also results in a transformation of vacuum into a thermal bath [13]. Since the concepts of particle number and vacuum play a central role in the interpretation of quantum field theory, the fact that they are not preserved in accelerated frames raises severe paradoxes and spoils the attempts to interpret the Einstein equivalence principle in the quantum domain [14].

With a conformal representation of accelerated frames in contrast, vacuum

and photon number are consistently defined for inertial and accelerated observers. For simplicity, we consider the simple case of a scalar massless field propagating along a single direction in a two-dimensional spacetime

$$\varphi(u) = \int_0^\infty \frac{d\omega}{2\pi}\sqrt{\frac{\hbar}{2\omega}}\left(a_\omega e^{-i\omega u} + a_\omega^\dagger e^{i\omega u}\right) \tag{2}$$

We use natural spacetime units ($c = 1$); ω represents the frequency as well as the wavevector; a_ω and $a_\omega^\dagger$ are the standard annihilation and creation operators. The total number of photons in the field state

$$N = \int_0^\infty \frac{d\omega}{2\pi} a_\omega^\dagger a_\omega \tag{3}$$

is invariant under conformal transformations to accelerated frames. To state more precisely this property, we introduce the conformal generators of quantum field theory as moments of the stress tensor [15]

$$\begin{aligned} E &= \int e(u)du \qquad & D &= \int ue(u)du \\ C &= \int u^2 e(u)du \qquad & e(u) &=: (\partial_u\varphi(u))^2 : \end{aligned} \tag{4}$$

The symbol : : prescribes a normal ordering of products of operators, so that the energy density $e(u)$ and consequently the conformal generators vanish in vacuum. The first generator E is the energy-momentum, that is also the translation operator associated with the light-cone variable u. Similarly, D corresponds to dilatations of this variable and C to transformations to accelerated frames [10]. Now the invariance of photon number under conformal transformations reads as

$$[E, N] = [D, N] = [C, N] = 0 \tag{5}$$

Hence, the action of any generator amounts to a redistribution of particles in the frequency domain without any change of the total particle number [10]. As a particular case, vacuum has the same definition ($N = 0$) for inertial and uniformly accelerated observers defined through conformal transformations [9]. For a massless field theory in two-dimensional spacetime, there exist higher-order conformal transformations which preserve field propagation, but not vacuum and photon number. This is consistent with the phenomenon of radiation from mirrors moving in vacuum with a non-uniform acceleration [16]. Here, we consider only the generators E, D and C which preserve N and, moreover, formally correspond to conformal transformations in four-dimensional spacetime.

We now give quantum definitions of the operators U and Ω associated with light-cone variable u and frequency ω. As the classical variables (1), the quantum references have to be preserved during propagation. They must therefore be built from the generators associated with conformal symmetry. We will simply define the operator U as the center of inertia of field energy, that is as the ratio of D and E (compare with eqs. (4))

$$U = D \cdot \frac{1}{E} \tag{6}$$

We take care of the non-commutativity of generators by introducing a symmetrised product represented by the $(\cdot)$ symbol. The definition (6) holds for any state orthogonal to vacuum ($E \neq 0$). It has this simple form because we consider a scalar field theory, or equivalently spin-0 particles. The operator Ω is then defined as the ratio of energy to particle number

$$\Omega = \frac{E}{\hbar N} \tag{7}$$

For a single particle state ($N = 1$), Ω plays exactly the same role as E. For a more general state however, the quantum fluctuations of Ω and E have independent meanings, since N also possesses its proper quantum fluctuations. Ω appears as a new quantum concept which represents the mean frequency of the field quanta and, like U, is defined for any state orthogonal to vacuum ($N \neq 0$). The expression needs not be symmetrized since E and N commute with each other.

Since the time and frequency references are defined from the conformal generators, their transformations are entirely determined by the commutators of these generators, that is by the conformal algebra

$$[E, D] = i\hbar E \qquad [E, C] = 2i\hbar D \qquad [D, C] = i\hbar C \tag{8}$$

Using only the first of these commutation relations, one deduces the shift of the operator U under frame transformations associated with E and D

$$[E, U] = i\hbar \qquad [D, U] = i\hbar U \tag{9}$$

These two commutators imply that U is shifted as the classical variable u in the translations and dilatations. In particular, the first one means that the operator U is canonically conjugated to the operator E. The frequency shifts are then obtained by combining the commutators of the conformal algebra with the invariance property (5) for the photon number

$$[E, \Omega] = 0 \qquad [D, \Omega] = -i\hbar\Omega \qquad [C, \Omega] = -2i\hbar\Omega \cdot U \tag{10}$$

As expected, the frequency-shift vanishes in a translation and is proportional to frequency in a dilatation, thus reproducing the Doppler shift of frequency under a Lorentz transformation. The third commutator represents a position dependent frequency shift arising from transformations to accelerated frames. It fits exactly the form of the classical Einstein effect [6], while being now a quantum statement valid for any state orthogonal to vacuum. To avoid any ambiguity, we recall that the operators U and Ω are, like the classical variables u and ω, conserved during propagation, and may thus be used to transfer time or frequency information between remote observers. However this does not imply that the operators are invariant under frame transformations. Clearly, these operators are shifted when different frames, that is also different observers, are considered. These shifts, which are described by commutation relations (9,10), reveal the basic relativistic properties of time and frequency, within the framework of quantum theory.

The time-shift under transformation to accelerated frames may also be obtained from the conformal algebra

$$[C, U] = i\hbar \left(U^2 - \frac{\alpha^2}{E^2} \right) \qquad \alpha^2 = C \cdot E - D^2 + \frac{\hbar^2}{4} \tag{11}$$

The quadratic form α^2 is a Casimir invariant of the conformal algebra (8) which commutes with all generators. Strikingly, it has a universal value $\alpha^2 = \frac{\hbar^2}{4}$ in any 1-particle state, independently of the frequency or time distribution of this state. The first term in the commutator $[C, U]$ corresponds to the classical shift. The second term appears as a correction associated with the pulse duration [5]. The transformation of U under C thus differs from its classical analog while the transformation of Ω coincides with its classical analog. It is however possible to state in a precise manner the consistency between time- and frequency-shifts in the quantum domain. As a matter of fact, the correction $\frac{\alpha^2}{E^2}$ to the time-shift (11) commutes with E, and is therefore invariant under a time translation. It is thus unchanged when the field pulse used for the transfer is delayed. This implies that clock rates, which correspond to differences between delayed clock indications, and frequencies vary consistently in transformations to accelerated frames, although the time shifts depart from the predictions of classical covariance rules.

A fact of great interest for the physical analysis of time-frequency transfer is that the shifts are now expressed in a theoretical framework where quantum fluctuations may be analyzed. The various commutators discussed above may

indeed be considered as setting the quantum limits in clock synchronisation. Since N commutes with all conformal generators, it also commutes with U and Ω, so that there is no constraint on number fluctuations in time-frequency transfer. Then time and frequency references obey an Heisenberg inequality

$$[\Omega, U] = \frac{i}{N} \qquad \Delta\Omega \cdot \Delta U \geq \frac{1}{2N} \tag{12}$$

In the limiting case of a large number of particles, it is possible to build nearly dispersionless field pulses and therefore to perform time-frequency transfer in a semiclassical regime. In the general case, the previously written relations provide a consistent description of the relativistic transformations and of the quantum fluctuations of time and frequency references.

References

[1] A. Einstein, Annalen der Physik **17** 891 (1905).

[2] See for example the Special Issue of the Proc. of IEEE **79** 891-1079 (1991).

[3] H. Salecker and E.P. Wigner, Phys. Rev. **109** 571 (1958).

[4] C. Rovelli, Class. Quantum Gravity **8** 297, 317 (1991).

[5] M.T. Jaekel and S. Reynaud, Phys. Rev. Lett. **vv** pppp (1996).

[6] A. Einstein, Jahrb. Rad. Elektr. **4** 411 (1907).

[7] H. Bateman, Proc. of London Math. Soc. **8** 223 (1909); E. Cunningham, ibidem **8** 77 (1909); T. Fulton, F. Rohrlich and L. Witten, Rev. Mod. Phys. **34** 442 (1962).

[8] E.L. Hill, Phys. Rev. **67** 358 (1945); **72** 143 (1947); T. Fulton, F. Rohrlich and L. Witten, Nuovo Cimento **26** 653 (1962).

[9] M.T. Jaekel and S. Reynaud, Quantum Semiclassical Opt. **7** 499 (1995).

[10] M.T. Jaekel and S. Reynaud, Brazilian J. of Phys. **25** 315 (1995).

[11] A. Einstein, Annalen der Physik **17** 132 (1905).

[12] W. Rindler, Essential Relativity (Springer, 2nd edition, 1977).

[13] P.C.W. Davies, J. of Phys. **A 8** 609 (1975); W.G. Unruh, Phys. Rev. **D 14** 870 (1976); N.D. Birrell and P.C.W. Davies, Quantum Fields in Curved Space (Cambridge, 1982).

[14] W.G. Unruh and R.M. Wald, Phys. Rev. **D 29** 1047 (1984); V.L. Ginzburg and V.P. Frolov, Sov. Phys. Usp. **30** 1073 (1987) [Usp. Fiz. Nauk **153** 633 (1987)].

[15] C. Itzykson and J.-B. Zuber, Quantum Field Theory (McGraw Hill, 1985).

[16] S.A. Fulling and P.C.W. Davies, Proc. Roy. Soc. **A 348** 393 (1976); A. Lambrecht, M.T. Jaekel and S. Reynaud, contribution to this book.

Synchronisation d'horloge avec des champs quantiques
Résumé:

La comparaison de temps par transfert de signaux lumineux joue un rôle central dans la théorie de la relativité. La nature quantique des champs utilisés dans une procédure de synchronisation d'horloge conduit à rediscuter les notions de temps et de fréquence ainsi que leurs transformations dans les changements de référentiel.

On utilise l'invariance dans les transformations conformes de la théorie des champs quantiques de masse nulle, de la définition du vide et du nombre de particules. Ces transformations permettent en particulier de représenter les référentiels uniformément accélérés.

On définit les références de temps ou de fréquence qui peuvent être encodées dans un pulse et transférées d'un observateur à un autre. Ces références sont des observables quantiques, exprimées en fonction des générateurs associés à la symétrie conforme. Elles constituent donc des notions très différentes des concepts classiques de temps ou de fréquence, et présentent des fluctuations quantiques compatibles avec la quatrième relation de Heisenberg.

Les relations de commutation caractéristiques de l'algèbre conforme permettent ensuite de déduire les transformations de ces observables dans un changement de référentiel. La transformation du temps lors du changement d'un référentiel inertiel vers un référentiel uniformément accéléré diffère de la prédiction classique associée aux propriétés de covariance, tout en respectant la compatibilité des transformations des références de temps et de fréquence.

MOTION INDUCED RADIATION FROM A VIBRATING CAVITY

Astrid Lambrecht[a], Marc Thierry Jaekel[b] and Serge Reynaud[c]
(a) Max-Planck-Institut für Quantenoptik,
Hans-Kopfermann-Str.1, D-85748 Garching, Germany
(b) Laboratoire de Physique Théorique, ENS, CNRS, UPS,
24 rue Lhomond, F-75231 Paris Cedex 05, France
(c) Laboratoire Kastler Brossel, ENS, CNRS, UPMC,
4 place Jussieu, F-75252 Paris Cedex 05, France

Vacuum field fluctuations exert radiation pressure forces on any scatterer placed in empty space. For two mirrors at rest in vacuum, the effect of vacuum radiation pressure is well known as the Casimir effect [1]. It has more recently been recognized that dynamical counterparts of this static force appear for moving scatterers. For some types of motion the field does not remain in the vacuum state, but photons are produced through non adiabatic processes [2]. Motion induced effects of vacuum radiation pressure do not require the presence of two mirrors but already exist for a single mirror moving in vacuum. In this case, the radiation reaction force is known to arise as soon as the mirror has a non-uniform acceleration [3]. The effects of radiation from a moving mirror and the associated radiation reaction force raise intriguing questions to the standard mechanical description of motion. It would therefore be very important to obtain experimental evidence of these dissipative processes associated with motion in vacuum. But vacuum radiation pressure scales as Planck's constant $\hbar$ and produces therefore only small mechanical perturbations for any macroscopic mirror. However, the number of photons emitted from an empty

high finesse cavity oscillating in vacuum, is the ratio of the radiated energy to the photon energy and thus scales as $\hbar^0$. This argument clearly supports a detection of optical rather than mechanical signatures of motion induced dissipation. A number of calculations has been devoted to the energy build-up inside a cavity with perfect mirrors [4]. However these calculations do not provide satisfactory answers to the previously discussed questions. They do not consider the photons radiated by the cavity since the latter is treated like a closed system. Even for the photons produced inside the cavity, the hypothesis of perfect mirrors amounts to disregard the important problem of finite lifetime of photons inside the cavity. In contrast, we study the configuration of a cavity built with partly transmitting mirrors. The cavity thus appears as an open system able to radiate into the free field vacuum. At the same time, the influence of the cavity finesse may be quantitatively evaluated.

For simplicity, we limit ourselves here to two-dimensional space-time calculations. As is well known from the analysis of squeezing experiments, the transverse structure of the cavity modes does not change appreciably the physical conclusions drawn from this simple model. Each transverse mode is correctly described by the two-dimensional model as soon as the size of the mirrors is larger than the spot size associated with the mode.

In the following, we first consider briefly the case of a single moving mirror and calculate the photon flux as well as the spectrum of the emitted radiation. To derive the radiation we use general arguments associated with scattering theory, without specific assumptions on the form of the interaction between mirror and field. This approach does not rely on a detailed microscopic analysis and is therefore applicable to any type of mirror as long as its internal dissipation is negligible. We also consider that the recoil of the mirror can be disregarded. The transformation from the incoming to the outgoing free field is then described by an S-matrix. We now introduce column-matrices $\Phi[\omega]$ which contain the components at a given frequency ω of the free fields which propagate in opposite directions

$$\Phi[\omega] = \sqrt{\frac{\hbar}{2|\omega|}} \begin{bmatrix} \left(\theta(\omega) a_{+,\omega} + \theta(-\omega) a^{\dagger}_{+,-\omega}\right) \\ \left(\theta(\omega) a_{-,\omega} + \theta(-\omega) a^{\dagger}_{-,-\omega}\right) \end{bmatrix} \tag{1}$$

Field components with positive or negative frequencies correspond respectively

to annihilation ($a_{\pm,\omega}$) and creation ($a^{\dagger}_{\pm,\omega}$) operators (θ is the Heaviside step function). The scattering of the field on a motionless mirror does not change the field frequency. The transformation from the input field $\Phi_{\rm in}$ to the output one $\Phi_{\rm out}$ is described by a unitary S-matrix which contains the transmission and reflection amplitudes $s\,[\omega]$ and $r\,[\omega]$ at a given frequency. Vacuum is preserved by a motionless mirror, as a consequence of unitarity.

When the mirror is moving, the frequency of the field is changed by the scattering process and the S-matrix now describes this frequency change

$$\Phi_{\rm out}\,[\omega] = \int \frac{d\omega'}{2\pi} S\,[\omega,\omega']\,\Phi_{\rm in}\,[\omega'] \tag{2}$$

Assuming that the incoming field is in the vacuum state, one obtains the following expression for the number N of photons radiated into vacuum by the moving mirror

$$\begin{aligned} N &= \int_0^{\infty} \frac{d\omega}{2\pi} \int_0^{\infty} \frac{d\omega'}{2\pi} n\,[\omega,\omega'] \\ n\,[\omega,\omega'] &= \frac{\omega}{\omega'} \mathrm{Tr}\left(S\,[\omega,-\omega']\,S\,[\omega,-\omega']^{\dagger}\right) \end{aligned} \tag{3}$$

$n\,[\omega,\omega']$ is the spectral density which describes the number of particles present in the output field. When the S-matrix is evaluated in a first order expansion in the displacement the spectral density $n\,[\omega,\omega']$ is proportional to the square modulus of the frequency component $\delta q\,[\omega+\omega']$ of the displacement

$$\begin{aligned} n\,[\omega,\omega'] &= \frac{\omega\omega'}{c^2} \gamma\,[\omega,\omega']\,|\delta q\,[\omega+\omega']|^2 \\ \gamma\,[\omega,\omega'] &= 2(1 - s\,[\omega]\,s\,[\omega'] + r\,[\omega]\,r\,[\omega'] \\ &\quad +1 - s\,[\omega]^*\,s\,[\omega']^* + r\,[\omega]^*\,r\,[\omega']^*) \end{aligned} \tag{4}$$

This expression results from a linear approximation of the motional perturbation of the field, but it is valid without any restriction on the motion's frequency.

We consider now the case of a mirror following a harmonic motion at a frequency Ω. Since we expect the radiation of photons to be proportional to time, we focus our attention onto a harmonic motion of amplitude a during a time T. For a long oscillation time T, we find the number N of radiated photons to be defined per unit time

$$\frac{N}{T} = \frac{a^2}{c^2} \int_0^{\Omega} \frac{d\omega}{2\pi} \omega\,(\Omega-\omega)\,\gamma\,[\omega,\Omega-\omega] \tag{5}$$

This result is similar to the expression one would obtain for the number of photons spontaneously emitted by an atom coupled to vacuum fluctuations, calculated with Fermi's golden rule. Here the emission is generated by the parametric coupling of the mirror's mechanical motion to vacuum radiation pressure rather than by the coupling of the atomic dipole to the vacuum field. Hence photons are emitted through a two-photon parametric process rather than through a one-photon process. As is well known, spontaneous emission is not accompanied by absorption processes because vacuum is the field ground state. Here the same property entails that photons are only emitted at frequencies ω and ω' smaller than the frequency Ω of the mechanical motion.

In the limiting case of a nearly perfect mirror ($s \to 0; r \to -1$), we obtain a simplified expression for the number of radiated photons

$$\begin{aligned} \frac{N}{T} &= \frac{8a^2}{c^2} \int_0^{\Omega} \frac{d\omega}{2\pi} \omega \, (\Omega - \omega) = \frac{2a^2\Omega^3}{3\pi c^2} \\ N &= \frac{\Omega T}{6\pi} \left(\frac{v}{c}\right)^2 \qquad v = 2\Omega a \end{aligned} \tag{6}$$

Expression (6) for N is a product of two dimensionless factors, namely the number of mechanical oscillation periods during the time T and the square of the maximal velocity v of the mirror divided by the velocity of light c. A characteristic feature of motion induced radiation, which could be used in an experiment to distinguish it from spurious effects, is the parabolic shape of its spectral density with a maximum at $\omega = \frac{\Omega}{2}$.

The derivation of motion induced radiation is similar in the case of two moving mirrors. Assuming the two mirrors to follow a harmonic motion at the same frequency Ω with respective amplitudes a_i $(i = 1, 2)$, we deduce the number of photons radiated per unit time to be

$$\frac{N}{T} = \sum_{ij} \frac{a_i a_j}{c^2} \int_0^{\Omega} \frac{d\omega}{2\pi} \omega \, (\Omega - \omega) \, \gamma_{ij} \, [\omega, \Omega - \omega] \tag{7}$$

As in the case of a single mirror, the functions γ_{ij} already appear in the evaluation of motional forces and they have been studied previously [6]. We introduce here simplifying assumptions allowing to obtain analytical expressions for the motional radiation. In the frequency range $[0, \Omega]$ one can in a good approximation assume the reflexion coefficients r_1 and r_2 of the two mirrors to be real and frequency independent. In the following we are concentrating on the most

interesting case where the cavity has a high finesse which implies that both r_1 and r_2 are close to unity. For this case we perform a resonant aprroximation, where the functions γ_{ij} only depend on the product $r_1 r_2$ of the reflection coefficients which we denote

$$r_1 r_2 = e^{-2\rho} \qquad \rho \ll 1 \tag{8}$$

where $\frac{1}{\rho}$ measures the cavity finesse. They then read

$$\begin{aligned}
&\gamma_{11}[\omega,\omega'] = \gamma_{22}[\omega,\omega'] = 4 + 4D_+[\omega]\, D_+[\omega'] \\
&\gamma_{12}[\omega,\omega'] = \gamma_{21}[\omega,\omega'] = -4D_-[\omega]\, D_-[\omega'] \\
&D_+[\omega] = \frac{\sinh(\rho)\cosh(\rho)}{\sinh^2(\rho) - \sin^2(\omega\tau)} = \sum_{k=-\infty}^{\infty} \frac{\rho}{\rho^2 + (\omega\tau - k\pi)^2} \\
&D_-[\omega] = \frac{\sinh(\rho)\cos(\omega\tau)}{\sinh^2(\rho) - \sin^2(\omega\tau)} = \sum_{k=-\infty}^{\infty} \frac{(-1)^k \rho}{\rho^2 + (\omega\tau - k\pi)^2}
\end{aligned} \tag{9}$$

τ is the time of flight of a photon from one mirror to the other. With exception of the first term in γ_{11}, all terms contain denominators clearly associated with the presence of the cavity. We can now calculate the emitted photon number by performing the integration (7) for the various Lorentzian components of the spectrum. Using the assumption of a high finesse cavity we find

$$\begin{aligned}
&\frac{N}{T} = \frac{\Omega^3}{3\pi c^2}(a_1^2 + a_2^2) + \frac{\Omega}{6\pi c^2}(\Omega^2 - \frac{\pi^2}{\tau^2})\left(\Delta_e(a_1 - a_2)^2 + \Delta_o(a_1 + a_2)^2\right) \\
&\Delta_e[\Omega] = \frac{\sinh(\rho)\cosh(\rho)}{\sinh^2(\rho) - \sin^2(\frac{\Omega\tau}{2})} = \sum_{k \text{ even}} \frac{4\rho}{4\rho^2 + (\Omega\tau - k\pi)^2} \\
&\Delta_o[\Omega] = \frac{\sinh(\rho)\cosh(\rho)}{\sinh^2(\rho) - \cos^2(\frac{\Omega\tau}{2})} = \sum_{k \text{ odd}} \frac{4\rho}{4\rho^2 + (\Omega\tau - k\pi)^2}
\end{aligned} \tag{10}$$

The first term in these expressions is a non-resonant contribution coming from direct reflection of vacuum fluctuations on both sides of the cavity. All other terms describe resonances of the motional radiation occuring when the mechanical excitation frequency Ω is close to an integer multiple of the fundamental optical resonance frequency $\frac{\pi}{\tau}$ and give rise to several emission peaks. Compared to the result obtained for a single mirror, the radiated photon flux is enhanced by a resonance factor which is essentially the cavity finesse. The emission peaks all have the same spectral width given by the cavity finesse and their relative intensities reproduce a parabolic spectrum, as the one obtained for a single moving mirror, however with a large resonant enhancement.

In equation (10), even modes $\Omega = \frac{2\pi}{\tau}, \frac{4\pi}{\tau} \ldots$ appear as elongation modes corresponding to a periodic modulation of the mechanical cavity length. More strikingly there also exist odd modes $\Omega = \frac{3\pi}{\tau}, \frac{5\pi}{\tau} \ldots$ which are excited by a global translation of the cavity with its length kept constant. This effect is reminiscent of radiation of a single oscillating mirror, since the cavity moves in vacuum without any further reference than vacuum itself. However radiation is now enhanced by the cavity finesse. These two kinds of vibration modes appear to be contrasted in a mechanical point of view, but not in the point of view of the field bouncing back and forth in the cavity. Indeed, the optical length as seen by the field varies similarly in both cases, although the mechanical cavity length is modulated in one case and constant in the other one.

So far we have given a quantitative estimate for the photon flux. The model of a cavity with partly transmitting mirrors allows us to evaluate now the resonance enhancement factor in terms of the cavity finesse. A mechanical excitation at exact resonance leads to the following orders of magnitude

$$N \simeq \frac{\Omega T}{2\pi} \frac{v^2}{c^2} \frac{1}{\rho} \tag{11}$$

where v measures the geometrical mean of the peak velocities of the two vibrating mirrors. To estimate the stationary number of photons inside the cavity, we may use a simple balance argument. Each photon has a probability 4ρ of escaping from the cavity during each roundtrip time 2τ. As we know the photon flux emitted by the cavity per unit time, we can deduce the number of photons $\mathcal{N}$ produced by the oscillation in a pair of cavity modes to be

$$\mathcal{N} \simeq \frac{v^2}{c^2} \frac{1}{\rho^2} \tag{12}$$

We may emphasize that not only the number of photons inside the cavity, but also the number of radiated photons, diverge at the limit of perfectly reflecting mirrors.

To be more specific about the orders of magnitude, let us recall that we have assumed the input fields to be in the vacuum state. This assumption requires the number of thermal photons per mode to be smaller than 1 in the frequency range of interest ($\hbar\omega < k_B\Theta$ with k_B the Boltzmann constant and Θ the temperature). Low temperature technology thus points to experiments

using small mechanical structures providing optical resonance frequencies as well as mechanical oscillation frequencies in the GHz range. In this frequency range, the finesse of a superconducting cavity can reach 10^9 [7]. A peak velocity $v \simeq 1$ m/s, corresponding to an amplitude in the nm range, would thus be sufficient to obtain a radiated flux of 10 photons per second outside and a stationary number of 10 photons inside the cavity. It is important to emphasize that the peak velocity considered in the present analysis is only a small fraction of the typical sound velocity in materials so that fundamental breaking limits do not oppose to these numbers. The photons may be detected, either by performing sensitive photon-counting detection of the radiated flux or by probing the state of the intra-cavity field using Rydberg atoms [7]. Therefore, if a technique is found to excite a vibrating motion with the above characteristics, the challenge of an experimental observation of motional radiation in vacuum can be taken up.

References

[1] H.B.G. Casimir, *Proc. K. Ned. Akad. Wet.* **51** 793 (1948); G. Plunien, B. Müller and W. Greiner, *Physics Reports* **134** 87 (1986)

[2] G.T. Moore, *J. Math. Phys.* **11** 2679 (1970)

[3] S.A. Fulling and P.C.W. Davies, *Proc. of the Royal Soc. of London* **A348** 393 (1976)

[4] C.K. Law *Phys. Rev. Lett.* **73** 1931 (1994); V.V. Dodonov, A.B. Klimov and V.I. Man'Ko, *Phys. Lett.* **A142** 511 (1989), **A149** 225 (1990); V.V. Dodonov, *Phys. Lett.* **A207** 126 (1995); see also references therein

[5] M.T. Jaekel and S. Reynaud, *Quantum Optics* **4** 39 (1992)

[6] M.T. Jaekel and S. Reynaud, *J. de Physique* **I-2** 149 (1992)

[7] G. Rempe, F. Schmidt-Kaler and H. Walther, *Phys. Rev. Lett.* **64** 2783 (1990); M. Brune, F. Schmidt-Kaler, A. Maali, J. Dreyer, E.Hagley, J.-M. Raimond and S. Haroche *Phys. Rev. Lett.* **76** 1800 (1996)

MEASUREMENT OF THE CASIMIR FORCE IN AN OPTICAL CAVITY WITH A RESONANT TRANSDUCTION SCHEME

P. Puppo, Univ. of Firenze, Dept. of Phys., P.le E. Fermi 2, I-50185 Firenze

Abstract

In 1948 Casimir and Polder[1], with their first calculations of the pressure between two conducting plates, showed the existence of the Casimir effect as a consequence of the vacuum fluctuations of the electromagnetic field. We present a experimental study to measure dynamically the Casimir force in a confocal optical cavity, using a mechanical resonant transduction scheme.

In a free cavity we cannot avoid to take into account the energy $\hbar\omega/2$ of each mode of the quantized electromagnetic field which is present even if no measurable photons are present. The modes and, therefore, the zero-point energy, directly related to the Heisenberg uncertainty principle, depend on the boundary conditions which are fixed by the geometry of the system. The vacuum energy is always infinite, but its change leads to the important physical consequence of the Casimir effect[2-4].

Several experiments were performed to measure the Casimir force[5]. In particular Sparnaay reported the test of the Casimir pressure:

$$P_c = \frac{1}{240}\pi^2\frac{\hbar c}{d^4} \tag{1}$$

using two conducting plates at distance $0.3 - 2.5\ \mu m$.

All these experiments had to confront with the smallness of the force and the difficulty to disentangle the retarded from the non retarded (Casimir force) part of the Van der Waals force. For instance, in the case of Sparnaay experiment, the data could fit with Casimir behaviour which decreases as d^{-4} as well as with the non-retarded force which decreases as d^{-3}, accordingly with the London theory.

Here we will present an experiment which permits to measure the Casimir force at macroscopic distances, in an confocal optical cavity by means of a resonant transducer.

In a lossles optical cavity formed by two spherical mirrors of curvature R at distance d, the allowed normal modes TEM_{pmq}, computed solving the self-consistent field equations are:

$$\omega_{pmq} = \frac{\pi c}{d}\left[q + \frac{1}{\pi}(2p + m + 1)\cos^{-1}\left(1 - \frac{d}{R}\right)\right] \tag{2}$$

where q labels the longitudinal modes and (pm) the transverse ones.

If the distance between the two mirrors is equal to R, i.e. in a confocal cavity, a degeneracy between longitudinal and tranverse mode appears: according to equation (2) all the modes $TEM_{k-j,2j,q-k}$ resonate at the same frequency ω_{00q} when $0 \leq j \leq k$.

Let us suppose that one of the cavity mirrors is perfectly reflecting at the frequency Ω within a bandwidth smaller than the distance between two subsequent longitudinal modes (free spectral range)

$$\Delta\Omega < \frac{\pi c}{d} \tag{3}$$

As a consequence, only the longitudinal mode ω_{00q} will resonate when $\Omega - \delta\Omega/2 \leq \omega_{00q} \leq \Omega + \delta\Omega/2$. This can be written as a resonance condition on the parameter q:

$$q \simeq \frac{d\Omega}{\pi c} \tag{4}$$

When the longitudinal mode q will resonate in the cavity, also the modes in degenaracy with it will do the same, therefore the zero-point energy will be:

$$E(d) = \frac{1}{2}\hbar\omega_{00q}N(\omega_{00q}) = E \tag{5}$$

where $N(\omega_{00q}) = (a^2/4d)^2(\omega/2\pi c)^2$ is the number of transverse modes resonating at the same frequency ω_{00q}, calculated taking into account the diffraction losses due to the finite size of the mirrors having diameter a. From condition (4) we can argue that varying the distance d between the two mirrors of a quantity

$$\delta = \frac{\pi c}{2\Omega} \tag{6}$$

it is possible to switch from the resonance configuration whose zero-point energy is E to the total absence of resonance with a null energy value. As a consequence a Casimir force arises between the two mirrors of the cavity, which is modulated with a period 2δ, and its magnitude can be estimated as[6]:

$$F_C = \frac{E}{\delta} = 4\pi\left(\frac{\hbar c}{\lambda_0^2}\right)\left(\frac{a^2}{4d\lambda_0}\right)^2 \tag{7}$$

where $\lambda_0 = 2\pi c/\Omega$ is the central cavity wavelength.

The choice of the geometrical parameter of the confocal cavity has to fulfill two important constraints.

The first condition is directly related to equation (3), which can be rewritten as

$$\frac{\Delta\Omega}{\Omega} < \frac{\pi c}{(d\Omega)} \tag{8}$$

The second constraint is concerning the confocality condition. If we vary the distance $d = R$, the mode degeneracy is indeed partially reduced. As a consequence F_C decreases. For this reason we have to evaluate how much we can move the mirrors, keeping the transverse

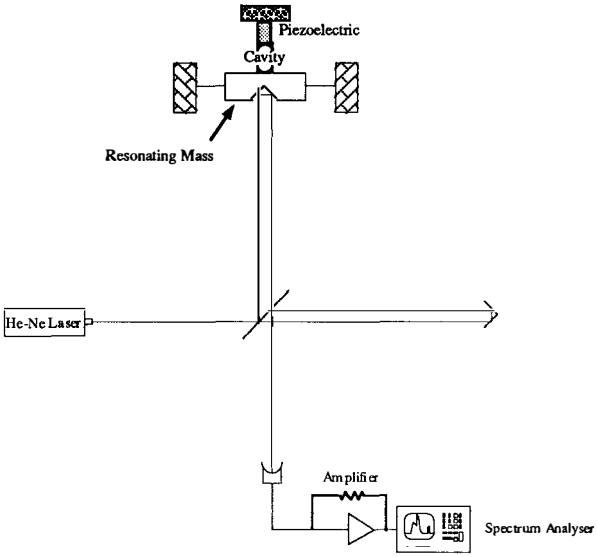


Figure 1: Experimental apparatus for the measurement of the Casimir force at long distances

modes degeneration. If $d = R(1 - \epsilon)$ it can be shown that the number of degenerate modes results bounded by the condition

$$N < (\pi/\epsilon)^2 \tag{9}$$

The experimental apparatus we are setting up is shown in the Fig. 1. In the optical cavity the mirrors have a diameter $a = 1\ cm$ and a curvature $R = 1.5\ cm$ and one of them is a notch optical filter with a bandwidth $\Delta\Omega \simeq 10^{-5}\Omega$ and $\lambda_0 = 1064nm$.

The variation of the distance d is perfomed using a cilindric piezoelectric actuator which moves one mirror with a maximum displacement $l = 20\ \mu m$ ($\epsilon = 1.3 \times 10^{-3}$, $N < 5.6 \times 10^6$). With these characteristic parameters we have a degeneracy number $N = 9 \times 10^4$ which permits to satisfy both constraints (8,9). The value of the corresponding force computed using the formula (7) is:

$$F_C \simeq 7 \text{ mdyne}$$

In order to increase the displacement induced by the force, one of the mirrors is mounted on one face of a resonant mass $m = 70$ g, having a frequency of about $\nu_R = 75\ Hz$ with a mechanical quality merit factor $Q = 1000$. To obtain a time modulation of the force at frequency ν_R, being the force space modulated with a period 2δ, it is sufficient to move the mirror by means of a piezoelectric with a velocity $v = \nu_r \lambda_0/2 = 1.9 \times 10^{-5}\ m/s$, i.e. with a frequency $\nu_p = v/l = 1\ Hz$. Any mechanical coupling between the piezoelectric and the resonator is consequently avoided.

The maximum displacement of the resonating mass due to the Casimir force is:

$$x_{max} = |X(\omega_R)| = |H(\omega_R)|\frac{F_C}{m} = \frac{F_C Q}{m\omega_r^2} = 4\ nm \tag{10}$$

where $X(\omega)$ in the Fourier transform of the monitored displacement $x(t)$, $H(\omega) = 1/(\omega_R^2 - \omega^2 + i\omega\omega_R/Q)$ is the trasfer function of the harmonic oscillator.

The other face of the resonator is equipped with the mirror belonging to an interferometer which is used to detect the signal. The arms of the interferometer are $L = 12\ cm$, and we plan to use a frequency stabilised He-Ne laser to illuminate it.

All the experimental apparatus is located in a vacuum chamber and the measurement is performed at room temperature.

The principal noise contributions on the signal output are the thermal noise of the resonating mass, the shot noise and the pressure radiation noise of laser light.

The thermal noise, due to the brownian motion of the harmonic oscillator at its resonant frequency ν_R is:

$$x_{br} = \sqrt{\frac{k_b T}{m\omega_R^2}} = 5.2 \times 10^{-13} m \tag{11}$$

The shot and the radiation pressure noise, which are both related to the discrete nature of light, can be expressed in terms of equivalent displacement on the resonant mass:

$$x_{shot} = \sqrt{\frac{\hbar c \lambda_L \omega_R}{4\pi P_{in} Q}} = 8.7 \times 10^{-16}\ m \tag{12}$$

$$x_{rad.pres.} = \frac{1}{m\omega_R^2}\sqrt{\frac{2\pi\hbar P_{in}}{c\lambda}} = 3.8 \times 10^{-24}\ m \tag{13}$$

The values, obtained with $P_{in} = 1\ mW$, are well below the signal we are going to detect, as a consequence, they are negligible.

The mechanical parts of the experimental apparatus have been designed in order to have all resonances of the structures higher than the monitored one.

Thanks to the macroscopic dimension of the optical cavity, it will be possible to perform a high sensitive measurement with the total absence of the Van der Waals forces.

I wish to thank Prof. E. Iacopini for his support and encouragement in the realisation of this experiment.

References

[1] H. B. G. Casimir, Proc. K. Ned. Akad. Wet., **51**, 793 (1948)
[2] T. H. Boyer, Ann. Phys. (N.Y.), **56**, 474 (1970)
[3] M. Fierz, Helv. Phys., **33**, 833 (1960)
[4] G. Plunien, B. Muller, and W. Greiner, Phys. Rep., **134**, 87 (1986)
[5] P. W. Milonni and Mei-Li Shih, Cont. Phys., **33**, 5 (1992)
[6] E. Iacopini, Phys. Rev. A, 48,1 (1993)

MESURE DE LA FORCE DE CASIMIR DANS UNE CAVITÉ OPTIQUE PAR UN SCHÉMA DE TRANSDUCTION RÉSONANT

Résumé

En 1948 Casimir et Polder démonsterènt l'existence de l'effect Casimir comme conséquence des fluctuations du vide du champ électromagnétique. Nous preséntons un étude expérimental pour mesurer dynamiquement la force de Casimir dans une cavité confocale en utilisant un schéma de transduction résonant.

DATA SECURITY AND QUANTUM MECHANICS

Artur Ekert and Chiara Macchiavello
Clarendon Laboratory, University of Oxford, Parks Road, Oxford OX1 3PU, U.K.

Abstract

Is there such a thing as a truly secret code? The ancient art of concealing information has in the past been matched by the ingenuity of code-breakers. But a combination of quantum physics and cryptography promises to dash the hopes of would-be eavesdroppers, perhaps for good. We provide a brief introduction to quantum cryptography and comment on the current experimental developments in the field.

1 What is wrong with classical cryptology ?

The purpose of cryptography is to transmit information in such a way that access to it is restricted entirely to the intended recipients. Originally the security of a cryptotext depended on the secrecy of the entire encrypting and decrypting procedures; today we use ciphers for which the algorithm for encrypting and decrypting could be revealed to anybody without compromising the security of a particular cryptogram. In such ciphers a set of specific parameters, called a *key*, is supplied together with the plaintext as an input to the encrypting algorithm, and together with the cryptogram as an input to the decrypting algorithm. The encrypting and decrypting algorithms are publicly announced; the security of the cryptogram depends entirely on the secrecy of the key, and this key must consist of any *randomly* chosen, sufficiently long string of bits.

Once the key is established, subsequent communication involves sending cryptograms over a public channel which is vulnerable to total passive eavesdropping (e.g. public announcement in mass-media). However in order to establish the key, two users, who share no secret information initially, must at a certain stage of communication use a reliable and a very secure channel. Since the interception is a set of measurements performed by the eavesdropper on this channel, however difficult this might be from a technological point of view, in principle any *classical* key distribution can always be passively monitored, without the legitimate users being aware that any eavesdropping has taken place.

Mathematicians have tried hard to solve the key distribution problem. The 1970s brought a clever mathematical discovery in the shape of 'public key' systems[1,2]. In these systems users do not need to agree on a secret key before they send the message. They work on the principle of a safe with two keys, one public key to lock it, and another private one to open it. Everyone has a key to lock the safe but only one person has a key that will open it again, so anyone can put a message in the safe but only one person can take it out. These systems exploit

the fact that certain mathematical operations are easier to do in one direction than the other. The systems avoid the key distribution problem but unfortunately their security depends on unproven mathematical assumptions, such as the difficulty of factoring large integers (RSA - the most popular public key cryptosystem gets its security from the difficulty of factoring large numbers[2]). This means that if and when mathematicians or computer scientists come up with fast and clever procedures for factoring large integers the whole privacy and discretion of public-key cryptosystems could vanish overnight. Indeed, recent work in quantum computation shows that quantum computers can factorize much faster than classical computers[3].

2 Quantum Cryptography

While classical cryptography employs various mathematical techniques to restrict eavesdroppers from learning the contents of encrypted messages, in quantum mechanics the information is protected by the laws of physics. In classical cryptography an absolute security of information cannot be guaranteed. The Heisenberg uncertainty principle and quantum entanglement can be exploited in a system of secure communication, often referred to as "quantum cryptography"[4]. Quantum cryptography provides means for two parties to exchange a enciphering key, consisting of a random sequence of binary numbers, over a private channel with complete security of communication.
There are basically two main types of quantum cryptosystems for the key distribution, these are:

1. Cryptosystems with encoding based on non-commuting observables or non-orthogonal quantum states[5,6].

2. Cryptosystems with encoding based on quantum entanglement[7].

Quantum cryptosystems based on non-commuting observables or non-orthogonal quantum states can be explained with the following simple example. The system includes a transmitter and a receiver. A sender may use the transmitter to send photons in one of four polarisations: 0, 45, 90, or 135 degrees. A recipient at the other end uses the receiver to measure the polarisation. According to the laws of quantum mechanics, the receiver can distinguish between rectilinear polarisations (0 and 90), or it can quickly be reconfigured to discriminate between diagonal polarisations (45 and 135); it can never, however, distinguish both types. The key distribution requires several steps. The sender sends photons with one of the four polarisations which are chosen at random. For each incoming photon, the receiver chooses at random the type of measurement: either the rectilinear type or the diagonal type. The receiver records the results of the measurements but keeps them secret. Subsequently the receiver publicly announces the type of measurement (but not the results) and the sender tells the receiver which measurements were of the correct type. The two parties (the sender and the receiver) keep all cases in which the receiver measurements were of the correct type. These cases are then translated into bits (1's and 0's) and thereby become the key.
An eavesdropper is bound to introduce errors to this transmission because she does not know in advance the type of polarisation of each photon and quantum mechanics does not allow her to acquire sharp values of two non-commuting observables (here rectilinear and diagonal polarisations). The most general eavesdropping strategy can be described as a unitary transformation

that entangles the transmitted state to an external system which is subsequently measured by the eavesdropper. The two legitimate users of the quantum channel test for eavesdropping by revealing a random subset of the key bits and checking (in public) the error rate. Although they cannot prevent eavesdropping, they will never be fooled by the eavesdropper because any, however subtle and sophisticated, effort to tap the channel will be detected.
In practice the legitimate users face the problem of distinguishing errors due to entanglement with an eavesdropper (caused by her measurements) from errors due to the innocent channel noise which cannot be avoided even in the eavesdropping-free environment. Several mathematical techniques, most notably the 'privacy amplification'[8], were developed to solve the problem but they do not cover the whole repertoir of possible eavesdropping strategies. Current research efforts are concentrated on specifying practical security criteria which will make the protocol described above operable.
Let us now describe briefly the entangled-based quantum cryptography; the basic idea is as follows. A sequence of correlated particle pairs is generated, with one member of each pair being detected by each party (for example, a pair of so-called Einstein-Podolsky-Rosen photons, whose polarisations are measured by the parties). An eavesdropper on this communication would have to detect a particle to read the signal, and retransmit it in order for his presence to remain unknown. Otherwise, in a more sophisticated eavesdropping strategy the correlated pairs undergo an entanglement with the eavesdropper's system, that is subsequently measured. However, both strategies destroy the quantum correlation of one particle of a pair with the other, and the two parties can easily verify whether this has been done (without revealing the results of their own measurements) by communication over an open channel.
In a noisy quantum channel the distributed pairs become entangled also with the environment. It has been shown recently that the entanglement-based cryptosystems can be supplemented with a procedure known as 'quantum privacy amplification' which allows to distill a fraction of the distributed pairs[9] whose entanglement with any outside system (including the eavesdropper) is arbitrarily low. The legitimate users can then use the distilled pairs for establishing the key - these pairs are not entangled with eavesdropper's measuring device and therefore the key is guaranteed to be perfectly secure.

3 Practicalities

A photon polarization measurement scheme has been used to make a working quantum key distribution system in a laboratory at the IBM Thomas J. Watson Research Center[5], which transmitted over the modest length of 30 cm at a rate of 10 bits/second. This experiment was set up to demonstrate the principles of quantum cryptography rather than to build a working prototype. Progress in quantum optics has resulted in new photon sources, new photodetectors, and better optical fibres; the components which have the potential for exhibiting the relevant quantum phenomena over much larger distances. For example, polarization based scheme has been successfully tested over a distance of 1 km[10], quantum entanglement has been tested over a distance of 4 km[11], and single-photon interference fringes have been produced in transmission over 10 km of fiber optic cable[12] (recent experiments in the laboratory have also shown that single-photon interference schemes with key generation rates of a few kilo bits per second are already feasible for propagation distances of about 50 km.[13]). Moreover, in more

recent experiments a polarisation-based quantum cryptographic channel has been successfully implemented in the laboratory using 26 km of coiled optical fibre and in the outside world, below Lake Geneva, operating with standard installed Swiss Telecom PTT fibres over a distance of 23 km[14]. This holds out a reasonable prospect for implementation of a secure key distribution system on a large local area network, transmitting at about 20k bits per second with present technology.

A.E. is supported by the Royal Society and C.M. by the European HCM Programme.

References

[1] W. Diffie and M.E. Hellman, IEEE Trans. Inf. Theory IT-22, 644 (1977).

[2] R. Rivest, A. Shamir, and L. Adleman, "On Digital Signatures and Public-Key Cryptosystems", MIT Laboratory for Computer Science, Technical Report, MIT/LCS/TR-212 (January 1979).

[3] P.W. Shor, "Algorithms for quantum computation," draft of April 18, 1994 manuscript, AT-T Bell Laboratories.

[4] C.H. Bennett, G. Brassard, and A.K. Ekert, "Quantum cryptography," Scientific American, October 1992, p. 50.

[5] S. Wiesner, SIGACT News **15**, 78 (1983); original manuscript written circa 1970. C.H. Bennett and G. Brassard, in "Proc. IEEE Int. Conference on Computers, Systems and Signal Processing", IEEE, New York,(1984). C.H. Bennett, F. Bessette, G. Brassard, L. Salvail, and J. Smolin, "Experimental quantum cryptography," J. Cryptology **5**, 3 (1992).

[6] C.H. Bennett, Phys. Rev. Lett. **68**, 3121 (1992).

[7] A.K. Ekert, Phys. Rev. Lett. **67**, 661 (1991). A.K. Ekert, J.G. Rarity, P.R. Tapster, and G.M. Palma, Phys. Rev. Lett. **69**, 1293 (1992).

[8] C.H. Bennett, G. Brassard, and J.-M. Robert, SIAM J. Comp. **17**, 210 (1988); C.H. Bennett, G. Brassard, C. Crépeau, and U.M. Maurer, "Generalized privacy amplification", IEEE Trans. on Information Theory, Vol. IT-41, no. 6, November 1995, in press.

[9] D. Deutsch, A. Ekert, R. Jozsa, C. Macchiavello, S. Popescu and A. Sanpera, *Security of quantum cryptography over noisy channels*, subm. to Phys. Rev. Lett.

[10] A. Muller, J. Breguet, and N. Gisin, Europhys. Lett. **23**, 383 (1993).

[11] P.R. Tapster, J.G. Rarity and P.C.M. Owens, Phys. Rev. Lett. **73**, 1923 (1994).

[12] P D. Townsend, J.G. Rarity, and P.R. Tapster, Electron. Lett. **29**, 1291 (1993).

[13] R.J. Hughes, D.M. Alde, P. Dyer, G.G. Luther, G.L. Morgan and M. Schauer, Contemporary Physics **36**, 149 (1995).

[14] A. Muller, H. Zbindem and N. Gisin, Nature **378**, 449 (1995).

DISSIPATION AND QUANTUM MECHANICS WITH TRAPPED IONS

J.F. Poyatos[(1)], J.I. Cirac,[(1)] and P. Zoller[(2)]

(1) Departamento de Fisica, Universidad de Castilla-La Mancha, 13071 Ciudad Real, SPAIN
(2) Institut für Theoretische Physik, Universität Innsbruck, Technikerstrasse 25, A-6020 Innsbruck, Austria.

Recent progress in trapping and laser cooling has allowed to hold a single ion in the lowest level of the 3D trapping potential[1)]. Although the original idea was to use this system in high resolution metrology, it is now recognized that trapped ions are also ideally suited for fundamental studies on Quantum Mechanics[2)]. Under a specific regime, trapped ions have been proposed as an alternative implementation of the well–known Jaynes–Cummings Model, usually found in the cavity QED context. This model describes the interaction of a two–level system with a harmonic oscillator. In the case of a single trapped ion, the dissipation in the oscillator (motion in a harmonic potential) is very small, and the coupling between the two–level system and the harmonic oscillator can be easily monitored just changing the laser intensity or frecuencies.

In this work, we will concentrate on the interaction of trapped ions with an environment. As it is well known, this coupling leads to dissipation and decoherence in quantum systems, which cannot be considered anymore as closed systems following unitary evolution. These processes depend critically on the way the system couples to the environment. We will show how to

design this coupling in a situation that is experimentally accessible with already developed technology. The system under consideration will be an ion confined in an electromagnetic trap, and the environment will be the vacuum modes of the electromagnetic field. We stress that the trapped ion is also a unique system to study the effects of decoherence in view of the recent experimental and theoretical advances, since: (i) on can prepare a variety of non-classical states[3--4]; (ii) one can perform tomographic measurements that is, one can measure the complete density operator of the system after any given interaction time to characterize the effects of decoherence[5]. Thus, the method we propose gives experimental access to physical situations that so far have been regarded as simple theoretical models to acquire a better understanding of the process of decoherence. It also opens new possibilities to study other more exotic situations that have not been studied yet since they were considered to be too far from realizable experiments.

In general, the master equation describing the markovian coupling between a harmonic oscillator and the environment is ($\hbar = 1$) [6]

$$\dot{\rho} = -i[H_0, \rho] + \gamma(2f\rho f^\dagger - f^\dagger f\rho - \rho f^\dagger f), \tag{1}$$

where ρ is the reduced density operator for the system after tracing over the states of the environment, H_0 is the ion Hamiltonian, and f operators acting on the ion's Hilbert space (here we have considered a single decay channel). The operator f and the parameter γ reflects the system–environment coupling. For a harmonic oscillator, f will be a function of the creation and annihilation operators a and $a^\dagger$, which are defined as usual

$$X = \sqrt{\frac{1}{2M\nu}}(a^\dagger + a), \qquad P = i\sqrt{\frac{M\nu}{2}}(a^\dagger - a), \tag{2}$$

where X and P are the position and momentum operators and M the particle's mass.

The solution and interpretation of this master equation can be formulated in the light of continuous measurement theory. According to this theory, the evolution of a state $|\Psi\rangle$ in a given experiment takes place in steps, which are composed of: (a) continuous evolution under an effective non-hermitian Hamiltonian $|\Psi(t)\rangle = e^{-iH_{\rm eff}t}|\Psi\rangle$, where $H_{\rm eff} = H_0 - i\gamma f^\dagger f$; (b) quantum jump according to $|\Psi(t+\delta t)\rangle = f|\Psi(t)\rangle$ (the operator f is thus called jump operator). The way decoherence takes place on the system depends on the particular form of the jump operator f, which in turn depend on the coupling between the system and the environment, as well as on the state of the latter. In particular, a state $|\Psi\rangle$ fulfilling $f|\Psi\rangle = 0$ will be unaffected

by decoherence, and therefore it can be considered as the steady state (at least for a particular set of initial conditions). One might find numerous examples in the literature considering jump operators, for which the process of decoherence is qualitatively different. Unfortunately, from the *experimental* point of view there are only few physical systems in which the process of decoherence can be studied. Moreover, in most of them the coupling with the environment has a fixed form which has precluded so far from testing interesting theoretical predictions. Our goal is to design different Markovian couplings between a system and an environment, which would permit a systematic study of decoherence experimentally. The coupling between the motion of the ion (the system) and the environment (vacuum modes of the electromagnetic field) takes place through the absorption of a laser photon and subsequent spontaneous emission. We will show that for low laser intensities and by choosing the laser frequencies one can select the effective coupling between the motion of the ion and the environment.

Let us consider a single ion moving in a one–dimensional harmonic potential and interacting with a laser in a standing wave configuration of frequency ω_L, close to the transition frequency ω_0 of two internal levels $|g\rangle$ and $|e\rangle$. Using standard methods in quantum optics, the master equation that describes this situation can be written in the general form

$$\dot{\rho} = -iH_{\rm eff}\rho + i\rho H_{\rm eff} + \mathcal{J}\rho, \tag{3}$$

with

$$H_{eff} = \nu a^\dagger a + \frac{1}{2}\omega_0\sigma_z + \frac{\Omega}{2}\sin[\eta(a+a^\dagger)+\phi](\sigma_+ e^{-i\omega t} + \sigma_- e^{i\omega t}) - i\frac{\Gamma}{2}|e\rangle\langle e|, \tag{4}$$

$$\mathcal{J}\rho = \Gamma\int_{-1}^{1} du N(u) e^{-i\eta u(a+a^\dagger)}\sigma_-\rho\sigma_+ e^{i\eta u(a+a^\dagger)} \tag{5}$$

where H_{eff} is the effective non–hermitian Hamiltonian and $\mathcal{J}$ the recycling superoperator. Here, $\sigma_+ = |e\rangle\langle g| = (\sigma_-)^\dagger$ and $\sigma_z = |e\rangle\langle e| - |g\rangle\langle g|$ are usual spin–$\frac{1}{2}$ operators describing the internal transition, ν is the trap frequency, Γ the spontaneous emission rate, and $\eta = 1/(2M\nu)^{(1/2)}$ the Lamb–Dicke parameter, with M the ion mass. In the expression for the jump superoperator, the exponentials are related to the photon recoil that takes place in each spontaneous emission process, and the integral takes into account the different angles at which that photon can be emitted, with a normalized dipole pattern $N(u)$. In the Hamiltonian describing the coupling with the lasers, Ω is the Rabi frequency, and ϕ characterizes the relative position of the trap center with respect to the node of the laser standing wave. Here we will

assume that either $\phi = 0$ (excitation at the node of the standing wave) or $\phi = \pi/2$ (excitation at the antinode).

Our goal now is to find physical situations in which this master equation can be approximated by the one given in (1), in such a way that by selecting the laser parameters (laser frequency and intensity) we can have different jump operators f. As a first step, we will symplify the coupling describing in the previous Hamiltonian. To do that, we move to a rotating frame defined by the unitary operator $\mathcal{U} = e^{-i(\nu a^\dagger a + \frac{1}{2}\omega_0 \sigma_z)t}$. Following Ref. [3] we assume that: (i) For excitation at the node ($\phi = 0$), $\delta = \omega_L - \omega_0 = (2k+1)\nu$ ($k = 0, \pm 1, \ldots$) (ii) For excitation at the antinode ($\phi = \pi/2$), $\delta = 2k\nu$ ($k = 0, \pm 1, \ldots$). In this rotating frame expanding the above master equation in terms of the Lamb–Dicke parameter ($\eta \ll 1$), assuming $\Gamma \ll \nu$, and together with low intensity considerations allows one to include in the coupling Hamiltonian only on–resonance terms (i.e. to perform rotating wave approximations), we get $H_{eff} = \frac{\Omega'}{2}(\sigma_+ f + f^\dagger \sigma_-)$, where both Ω' and the form of the operator f depend on the frequency of the laser. For example, for $\delta = -\nu$, we have $f = a$, and $\Omega' = \Omega\eta/2$, whereas for $\delta = -2\nu$, $f = a^2$, and $\Omega' = -\Omega\eta^2/6$ (see Fig. 1). Apart from the strong confinement, in the first case, the RWA can be performed for $\Omega' \ll \nu$, whereas in the second case it is needed $\Omega^2/\nu \ll \Omega'$. This two conditions can always be fulfilled for low enough laser intensity, and together with $\Omega' \ll \Gamma$ define the low intensity limit. Although here we will be interested in jump operators at most quadratic on a and $a^\dagger$, it is clear that the above procedure can be easily generalized to include operators with higher orders.

In the next step we eliminate adiabatically the internal excited state using standard procedures of quantum optics. Physically, since $\Omega' \ll \Gamma$ the ion practically spends no time in the excited level and therefore we can elliminate it. Finally, expanding the resulting equation in powers of η, we find the desired master equation (1), with corrections of the order η^2. The master equation will be valid for times such that these corrections are not important, that is for times $t \ll (\gamma\eta^2\bar{n})^{-1}$ where $\bar{n}$ is the typical phonon number of the state of the ion. Nevertheless in the Lamb–Dicke limit this time can be much longer than the time requiered to reach the steady state using the approximated master equation.

According to our analysis, by varying the laser frequency we obtain the master equation with different jump operators f. Note that through the adiabatic elimination we are coupling effectively the motion of the ion with the environment. The fact that this coupling takes place

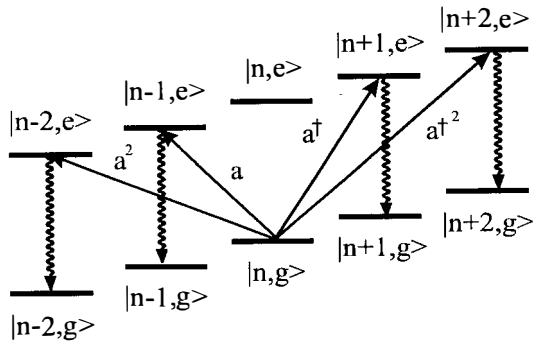


Figure 1: Level scheme for different laser configurations leading to various jump operators.

through the absorption of laser photons, and we have to choose the way on how this actually happens, allows to manipulate the coupling system–environment.

Let us show how to design a variety of jump operators. For example, consider the case in which two lasers of frequency $\omega_l + \nu$ and $\omega_L - \nu$ interact with the ion. In this case, following the same arguments as before, one can easily show that the jump operator is $f = \mu a + \nu a^\dagger$, where $\mu^2 - \nu^2 = 1$ and μ/ν is the quotient of the Rabi frequencies. Thus, one can produce a squeeze vacuum coupling[6], and vary the squeezing parameter at will. In particular, choosing equal Rabi frequencies, the jump operator is $f = a + a^\dagger \propto x$, which correspond to the case analized theoretically by Zurek[7], Caldeira and Legget[8] and other authors, to describe the dynamics of the decoherence process. Another interesting combination of lasers, give rise to the jump operator $f = a^2 - \alpha^2$, where α is a given c–number. In this case, one needs a weak laser on resonace and another one such that $\omega_L = \omega_0 - 2\nu$. The value of α^2 is determined by the quotient of the effective Rabi frequency corresponding to these laser excitation. There are other possibilities to design other jump operators. For example, it is possible to use the three spatial dimensions of the ion motion, and to excite the ion with lasers propagating along different directions to vary not only the Rabi frequency but also the Lamb–Dicke parameter.

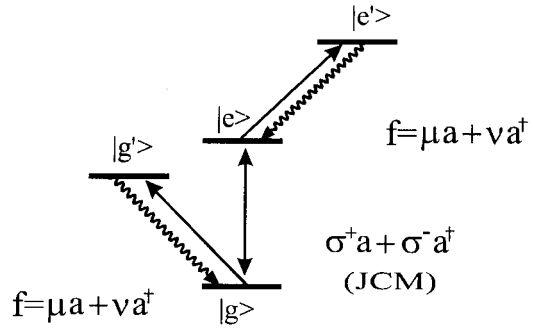


Figure 2: Internal level excitation corresponding to a two-level system in a squeezed vacuum.

One can also have a master equation with more than one jump operator by exciting two or more different internal transitions with lasers. There is also the possibility of having lasers coupling two different internal levels separately, so that one has a two-levels system coupled to a harmonic oscillator (the motion) which is damped . An interesting case is the one in which the damping of the harmonic oscillator corresponds to a squeezed reservoir, since in this case one can test experimentally some of the predictions made some time ago[6] about the decoherence in this problem (Fig. 2).

In summary, we have proposed a way to design master equations for experimental test of the ubiquous phenomena of decoherence in quantum mechanics. This opens a new testground to experience with many phenomena that so far have been only addressed theoretical point of view. With the recent set of experiments carried out recently by the group of Wineland at Nist, it is clear that our proposal is at reach of present technology.

REFERENCES:

1.- C. Monroe, D. M. Meekhof, B. E. King, S. R. Jefferts, W. M. Itano, D. J. Wineland, and P. Gould, Phys. Rev. Lett. **75**, 4011 (1995).

2.- See, for example, J. I. Cirac, A. S. Parkins, R. Blatt, and P. Zoller in Adv. in Atomic and Mol. Phys. (in press).

3.- J. I. Cirac, A. S. Parkins, R. Blatt, and P. Zoller, Phys. Rev. Lett. **70**, 556 (1993); J. I. Cirac, R. Blatt, A. S. Parkins, and P. Zoller, Phys. Rev. Lett. **70**, 762 (1993); The way to preparing arbitrary states of motion is shown in S. A. Gardiner, J. I. Cirac and P. Zoller (unpublished).

4.- C. Monroe, D. M. Meekhof, B. E. King, W. M. Itano, and D. J. Wineland, Phys. Rev. Lett. **75**, 4714 (1995); D. M. Meekhof, C. Monroe, B. E. King, W. Itano, and D. J. Wineland, (unpublished); C. Monroe, D. M. Meekhof, B. E. King, and D. J. Wineland, (unpublished).

5.- J. F. Poyatos, R. Walser, J. I. Cirac, P. Zoller, and R. Blatt, Phys. Rev. A (to be published); S. Wallentowitz and V. Vogel, Phys. Rev. Lett. **75**, 2932 (1995).

6.- See, for example, C. W. Gardiner, *Quantum Noise*, (Springer–Verlag, Berlin, 1992) and references therein.

7.- W. H. Zurek, Phys. Today **44**, No. 10, 36 (1991).

8.- A. O. Caldeira and A. J. Leggett, Phys. Rev. A **31**, 1059 (1985).

QUANTUM COMPUTING: DREAM OR NIGHTMARE

J.M. Raimond and S. Haroche
Laboratoire Kastler Brossel, Ecole Normale Supérieure
24 rue Lhomond, 75005, Paris, France

Abstract

A quantum computer is a theorist's dream: it operates exponentially faster than any ordinary one, making difficult problems (such as factorization) tractable. At the same time, it is an experimenter's nightmare since any large scale realization is doomed by the decoherence process. We discuss briefly this topic in the light of recent quantum optics experiments.

Present computer rely on classical physics to code and manipulate information. The elementary information is a bit, taking either 0 or 1 value. The calculation is an irreversible process driving the computer from its inital state (containing one possible input to the program) to its final state containing the corresponding result. A computer which would manipulate quantum systems (qbits) as the elementary information carriers, through a unitary quantum evolution[1, 2], would behave in a strikingly different way. Taking benefit of the superposition principle, it could process simultaneously all the possible inputs. This "massive quantum parallelism" enables the quantum computer to perform in a single run 2^n calculations on an n qbits input.

The interest in quantum computers raised recently with the discovery by Shor of an algorithm performing efficiently large integers factorization on a quantum computer[3, 4]. This problem is well known to be complex for standard computers. The best known algorithms[5] take a time growing almost exponentially versus size (number of bits), preventing numbers with more than 500 bits to be factorized in practical times. Large numbers factorization lies therefore at the heart of most cryptographic systems. Since the Shor's algorithm performs this factorization in a time growing only as the square of the input size, a quantum computer could break easily all the cryptograms.

It is not possible to give here a detailed account of this algorithm. We only emphasize the most important points. Like in a classical computer, the numbers are encoded into words of

n qbits, spanning a Hilbert space with dimension 2^n. A convenient basis (the so-called "computation basis") is the direct product of the individual qbits "logical levels" $|0\rangle$ and $|1\rangle$. Each of the 2^n kets in this basis can be represented as a string of binary digits, each corresponding to a single qbit, and associated to an integer number between 0 and $2^n - 1$. A calculation is a unitary transform changing each ket in the basis into another one. The Shor's algorithm makes use of two quantum registers. The values coded by one (and thus its state) depend upon the values (or upon the state) of the other. The two registers are said to be entangled: they are in a state which cannot be cast as a product of two independent ones. For large n's, such entangled states are macroscopic versions of the entangled spin pairs at the heart of the EPR paradox[6]. Quantum nonlocality therefore plays an important role in quantum computing: the projection postulate allows to alter one register's contents by measuring the other. This process is used in Shor's algorithm to extract the result of the computation. The central role of quantum projection makes the quantum computer very different from any optical computer, which could also be able to manipulate linear superpositions (of polarizations).

A quantum computation amounts in a series of unitary transforms performed on the entangled registers. Each transformation can be split into elementary operations involving only two or three qbits. By analogy with classical computers architecture, these elementary processes are called quantum gates or qgates[7, 8, 9]. Various types of qgates have been proposed. Most of them are "universal", meaning that any computation could be performed by a network made of interconnected copies of the same gate. Most of the two-bits qgates implement conditional quantum dynamics. The state of one of the qbits (the "control bit") controls whether or not the other (the "target") undergoes a unitary evolution. The control bit state is not altered. The simplest version is the "'controlled not" qgate[10]. When the control bit is in $|0\rangle$, the target bit is unchanged. When the control bit is in state $|1\rangle$, the target bit undergoes a π rotation, amounting to the transform $|0\rangle \rightarrow |1\rangle$ and $|1\rangle \rightarrow -|0\rangle$. This conditional dynamics clearly entangles the two qbits. This qgate can also be viewed as a measurement of the control qbit: if the target is initially in $|0\rangle$, its output state will be identical to the one of the control. This is a Quantum non Demolition process[11], since the state of the measured system is unchanged.

Various physical implementations of qgates have been proposed. The schemes for QND measurement at the single quantum limit can be transposed in a controlled not gate. For instance, we have proposed a QND field intensity measurement at the photon level using "circular" Rydberg atoms interacting with very high Q superconducting millimeter wave cavities[12]. The realization of this scheme would also demonstrate operation of a controlled not. Conditional dynamics has already been demonstrated experimentally[13], but the entanglement between the two qbits has not yet been observed. Improvements in progress now should allow us to explore this entanglement. Recent demonstration of similar conditional dynamics in an optical cavity quantum electrodynamics experiment has been performed[14]. Another very promising system is made of a single trapped, laser cooled ion. Laser manipulation of the internal and vibration degrees of freedom allows one to realize various types of conditional dynamics, including the controlled not gate as demonstrated recently[15]. The entanglement of the two qbits has not been evidenced, but it is clear that a full demonstration of a two bits quatum gate will occur in the next months.

Realization of a single gate does not mean that a large scale quantum computer could be built. At variance with ordinary computers, a quantum one is much more than the sum of its parts. The calculation succeeds only if the entanglement between all the involved qbits survives during the whole process. A macroscopic quantum computer would be at any stage of the computation in a superposition of many macroscopically different states. It would be even more complex than the cat imagined by Schrödinger to be in a quantum superposition of its dead and alive states[16]. Such macroscopic superpositions are extremely fragile, explaining

why they are never observed in our macroscopic world. The efficient coupling of large objects to thermodynamic reservoirs causes a fast evolution, called decoherence[17], from a quantum superposition to a statistical, incoherent mixture. The decoherence rate is proportional to the "size" of the system.

The simplest part of a quantum computer, an n qbits register, looses its coherence with a rate $n\Gamma$, where Γ is the decay rate for a single qbit. The calculation success probability is thus $\exp(-n\Gamma T)$, where T is the total calculation time, product of the number of elementary gate operations, t_{op}, multiplied by the number of operations, N_{op}. The maximum value of $N_{op}n$ is therefore of the order of the qbit "figure of merit", $1/(\Gamma t_{op})$. For most systems known as good candidates to the realization of a qgate, the ultimate figure of merit is between 10^4 and 10^6. Since factorization of a p bits number with the Shor's algorithm requires about $n = 5p$ storage bits and $N_{op} \simeq 300p^3$ operations[18], we get $nN_{op} \simeq 1500p^4$. A figure of merit of 10^6, achievable with present technology at the expense of an experimenter's nightmare, makes it possible to factor a $p = 4$ or 5 bits number, not quite a feat! Factoring an "interesting" number, with 600 bits, requires improving the figures of merit above this ultimate limit by at least 8 orders of magnitude!

One might think that the above limitations are of a mere technical nature, and could be overcome with the development of technology. We think it is quite not the case, and that there are fundamental limits to the figure of merit. Since we have to get an handle on the qbits states, in order to manipulate them during the qgates operations, this input channel necessarily adds some noise, at least of a quantum nature. If we consider, for instance, qbits coded in a two-level atomic system, such as in the cavity QED experiments, the noise is spontaneous emission, whose rate increases with the atomic dipole. Reduced noise means therefore a small dipole, and hence a long time to perform a system transformation by acting on the same dipole. This argument sets in fact a limit to the figure of merit of the order of $N\alpha^{-3/2}$, where N is the typical quantum number of the levels and α the fine structure constant. With $N = 50$, the limit value is of the order of 10^5, quite close to the already achieved one, and far from the requirements for a large scale computer. Let us mention that N cannot be arbitrarily increased, since the levels are increasingly sensitive to decoherence by stray electric field (the Stark polarisability of Rydberg levels scales like N^7!). In the case of the trapped ions technique, the situation is somewhat different, since the qbits manipulations involve Raman transitions between ground state levels, with basically infinite lifetime. However, spontaneous emission after an unwanted transition to the excited relay level also puts a limit to the figure of merit, of the order of α^{-3}, about 2.10^6. It is very likely that other system would lead to similar fundamental limits.

Decoherence puts very strong limitations on the feasibility of a quantum computer. The only escape would be to devise "error correction codes", restoring the quantum coherence in spite of the relaxation processes. Some schemes have been proposed[19], which could improve the effective figure of merit of the calculation by correcting errors at the first order. However, these codes add a tremendous overhead to the complexity of the system. As for classical error correction codes, they encode the information in a redundant way, using at least 3 qbits instead of a single one. Keeping in mind that, in spite of continued efforts, quantum physicists have not yet been able to produce a single triplet of entangled particules, it seems unlikely that they could be close to use such entangled states to process information. The time scales involved in the development of such devices with the present techniques is far beyond a man's life.

Quantum computing is a graal which will never be reached, unless a completely new technology appears. If it would be the case, it would have much farther reaching consequences than the mere calculation problem. It would indeed change completely our perception of the border between the classical and quantum worlds, since it would allow macroscopic objects to behave in a quantum way. Even if this is not possible, there are very interesting fundamental quantum

mechanics problems related to this quest. The first one is to get a better understanding, both theoretical and experimental, of complex entangled quantum systems. Realizing systems with three or four entangled qbits, testing the non locality of quantum mechanics, using them for new quantum cryptography schemes, would be a major breakthrough. The quantum computing vocabulary gives us valuable insights into the ways to prepare and manipulate such quantum systems. Decoherence is another major point. Besides dooming any attempt to build a large scale computer, this process plays a very important role, preventing our everyday life from being quantum. For all non elementary systems known today, the decoherence times are so short that decoherence has never been observed. Preparing a mesoscopic quantum coherence and observing it while it relaxes to an incoherent superposition is one of the most challenging issues of modern quantum physics. Finally, the study of quantum error correction codes is also very interesting for a next step. Their inherent complexity make them of little help to save large scale quantum computers, but the possibility to act on decoherence and to limit its effects to some extent is certainly a fascinating one.

In conclusion, it is important to make clear that, unless some completely new physics appears, the use of quantum mechanics to perform calculations is not a reasonable prospect. However, very important fundamental quantum mechanics problems at the heart of the "computing" process have a very large interest. Moreover, they can be explored experimentally in the simplest configurations (a few qbits) with cavity quantum electrodynamics or ion traps techniques. These problems are by themselves of such an importance that it is not necessary to oversell them as new schemes for computing, which would sooner or later backfire and cast serious doubts about the whole topic.

References

1) R. Feynman, Optics News **11**, 11 (1985).
2) D. Deutsch, Proc. Roy. Soc. London **A400**, 97 (1985).
3) P. Shor, *Proceedings of the 35th Annual Symposium on the Foundations of Computer Science* (IEEE Computer Society Press, Los Alamitos, CA, 1994), p. 124.
4) A. Ekert and R. Josza, Review of Modern Physics To be Published (1996).
5) *Lecture notes in Matehematics*, edited by A. Lenstra and H. Lenstra (Springer, Berlin, 1993), Vol. 1554.
6) A. Einstein, B. Podolski, and N. Rosen, Phys. Rev. **47**, 777 (1935).
7) D. Deutsch, Proc. Roy. Soc. London **A425**, 73 (1989).
8) D. di Vicenzo, Phys. Rev. **A51**, 1015 (1995).
9) D. Deutsch, A. Barenco, and A. Ekert, Proc. Roy. Soc. London **A449**, 669 (1995).
10) A. Barenco, D. Deutsch, A. Ekert, and R. Jozsa, Phys. Rev. Lett. **74**, 4083 (1995).
11) V. Braginsky and F. Khalili, Sov. Phys. JETP **46**, 705 (1977).
12) M. Brune *et al.*, Phys. Rev. Lett. **65**, 976 (1990).
13) M. Brune *et al.*, Phys. Rev. Lett **72**, 3339 (1994).
14) Q. Turchette *et al.*, Phys. Rev. Lett. **75**, 4710 (1995).
15) C. Monroe *et al.*, Phys. Rev. Lett. **75**, 4714 (1995).
16) E. Schrödinger, Naturwissenschaften **23**, 807,823,844 (1935).
17) W. Zurek, Phys. Today 36 (1991).
18) V. Vedral, A. Barenco, and A. Ekert, Preprint (1995).
19) P. Shor, Phys. Rev. **A52**, R2493 (1995).

EXPERIMENTAL GRAVITATION

VIRGO - A LASER INTERFEROMETER FOR THE DETECTION OF GRAVITATIONAL WAVES

Henrich Heitmann
on behalf of the VIRGO collaboration:

B. Caron, A. Dominjon, C. Drezen, R. Flaminio, X. Grave,
F. Marion, L. Massonnet, C. Mehmel, R. Morand, B. Mours, M. Yvert
LAPP, Chemin de Bellevue, B.P. 110, F-74941 Annecy-Le-Vieux Cedex

D. Babusci, S. Bellucci, G. Candusso, G. Giordano, G. Matone
INFN Frascati, Via E. Fermi 40, CP 13, I-00044 Frascati

L. Dognin, J.M. Mackowski, M. Napolitano, L. Pinard
IPN Lyon, Univ. C. Bernard, Lab. de Phys. Nuc., 43 Bd. du 11 Nov. 1918, F-69622 Villeurbanne Cedex

F. Barone, E. Calloni, L. Di Fiore, A. Grado, L. Milano, G. Russo, S. Solimeno
INFN Napoli, Univ. di Napoli, Dipt. di Scienze Fisiche, Pad. 19 Mostra d'Oltremare, I-80125 Napoli

M. Barsuglia, V. Brisson, F. Cavalier, M. Davier, P. Hello, F. LeDiberder, P. Marin, M. Taubman
Lab. de l'Accélérateur Linéaire, IN2P3-CNRS et Univ. de Paris-Sud, F-91405 Orsay

F. Bondu, A. Brillet, F. Cleva, H. Heitmann, L. Latrach, C.N. Man, Pham-Tu Manh, J.-Y. Vinet
VIRGO laser group, Bât 208 - Campus d'Orsay, F-91405 Orsay Cedex

C. Boccara, Ph. Gleyzes, V. Loriette, J.P. Roger
ESPCI (Laboratoire de Spectroscopie en Lumière Polarisée), 10 rue Vauquelin, F-75005 Paris

G. Cagnoli, L. Gammaitoni, J. Kovalik, F. Marchesoni, M. Punturo
INFN Perugia, Univ. di Perugia, Via A. Pascoli, I-06100 Perugia

M. Bernardini, S. Braccini, C. Bradaschia, R. Del Fabbro, R. DeSalvo, A. Di Virgilio, I. Ferrante,
F. Fidecaro, A. Gennai, A. Giassi, A. Giazotto, L. Holloway, P. La Penna, G. Losurdo,
F. Palla, Pan Hui-Bao, A. Pasqualetti, D. Passuello, R. Poggiani, G. Torelli, Zhang Zhou
INFN Pisa, Via Livornese 582/a, S.Piero a Grado, I-56010 Pisa

E. Majorana, P. Puppo, P. Rapagnani, F. Ricci
INFN Roma, Dipartimento di Fisica, Univ. La Sapienza, Piazza Aldo Moro 2, I-00185 Roma

Abstract

The French-Italian gravitational wave detector VIRGO[1-3] is presently under construction. It will be a Michelson type laser interferometer with 3 km long arms containing Fabry-Perot cavities for increasing the effective arm length, and using the technique of light recycling for increasing the effective laser power. This paper gives a short overview over the most important noise sources limiting the sensitivity of interferometric gravitational wave detectors, and describes the principal elements constituting the VIRGO interferometer.

Introduction

Gravitational waves

Gravitational waves are fluctuations of the geometry of space-time, predicted by Einstein's theory of General Relativity. The effect of a gravitational wave on two free floating particles sitting on given space-time coordinate points is a change in distance, like between two balls swimming on wavy water, if their distance is measured along the water surface. Gravitational waves (GW) are quadrupolar (i.e. tensorial), such that a ring of free particles is deformed to an expanding and contracting ellipse; accordingly, a mass distribution with a changing quadrupole moment emits GW. For example, the earth-moon system radiates 5.6 μW of GW energy, the earth-sun system 187 W, and two black holes of one solar mass turning around each other at 3 km distance 10^{50} W. From these values it is plausible, that laboratory generation of reasonably strong GWs is not possible; for a direct proof of their existence, GWs from astrophysical events are needed. Examples are:

- extremely close compact binary stars, losing energy due to GW and therefore inspiralling and finally colliding
- non-spherically symmetric supernova explosions
- rapidly rotating neutron stars with a quadrupole moment
- a stochastic GW background from indistinguishable individual sources or surviving from the big bang

The field strength of GW's is defined as the relative change in distance, $h = 2 \cdot \Delta L/L$ (dimensionless strain amplitude). Even if the transported energy is considerable, the resulting strain is extremely small: In the black hole example from above, the flux received on earth at a distance of 10 Mpc would be 100 W/m^2, but the induced relative length change h is only of the order of 10^{-21}. Therefore, extremely sensitive detection mechanisms are required.

Detection of Gravitational waves

The first experiments for the detection of GWs by Weber in the 1960's[4] were aluminum cylinders, whose internal resonances can be excited by a GW and read out capacitively. A different scheme uses the Michelson interferometer for detection. An incident GW, if properly polarized, increases one arm length and reduces the other one, and the resulting arm length difference can be detected at the interferometer output. Since no internal resonances are involved, a large detection bandwidth is possible (in the case of VIRGO, 10 Hz to 10 kHz). The sensitivity can be enhanced by taking longer arms, because the length difference ΔL is proportional to the arm length L. If the arm length exceeds one fourth of the GW wavelength, however, the signal reverses its sign before the light round-trip is completed, and the sensitivity decreases. For a typical GW frequency of 1 kHz, one obtains an optimum of 75 km, which is obviously very difficult to realize. But one can have the light perform several roundtrips in the arm before coming back to the beam splitter (optical delay line)[5], thus increasing the effective arm length. The alternative chosen by VIRGO consists in inserting an additional mirror in each arm, thus creating Fabry-Perot cavities in the arms which store the light[6]. In this way, the physical arm length can be reduced to 3 km, which is feasible, but still a

challenge. A further reduction of arm length is not possible because of the thermal noise of the mirrors (see below).

In order to minimize photon noise, the power of the laser illuminating the interferometer must be as high as possible. Since the detection sensitivity is optimum when the interferometer is operated at a dark fringe, almost all the light entering the interferometer (minus the losses) is reflected back towards the laser. By inserting an appropriate mirror, this light can be sent back to the interferometer (recycling)[7]; effectively, a resonant cavity is formed with a resonant enhancement which in the case of VIRGO is of the order of 50. Thus an input power of 20 W gives an effective power incident on the beam splitter of 1 kW.

The optical structure of VIRGO, resulting from these considerations, is shown in Fig. 1. The interferometer consists of one beam splitter and five mirrors.

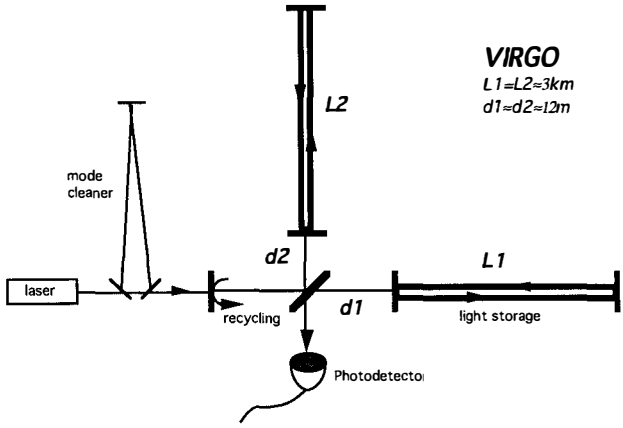


Fig. 1: Optical structure of the VIRGO interferometer.

The VIRGO interferometer and its main noise sources

The VIRGO interferometer is presently under construction in Italy with the participation of laboratories from the French CNRS (Annecy, Lyon, Orsay, Paris) and the Italian INFN (Frascati, Napoli, Perugia, Pisa, Roma). As it could already be seen in the introduction, the design of laser interferometric GW detectors is largely determined by the noise sources one wants to minimize. The following overview over the various noises affecting the detection sensitivity therefore allows to introduce some of the main aspects of VIRGO.

Seismic noise: Since the mirrors are not, as they would be in the ideal case, freely falling masses, unavoidable vibrations of the ground due to seismic human activity couple to them. All important optical components are therefore suspended as pendulums. A simple pendulum damps vibrations of the suspension point with $1/\nu^2$ above the resonance frequency ν_0. In VIRGO, the mirrors are suspended on seven-stage pendulums (superattenuator), where each stage provides seismic isolation in

all six degrees of freedom. A superattenuator thus gives a seismic isolation of $1/\nu^{14}$ above the resonance frequency, assuring that seismic noise becomes negligible above 3 Hz. In order to reduce the residual motion of the suspended mirror, the second stage is equipped with an electronic damping system which kills the most important resonances. In a first prototype, residual rms movements of a few μm were measured above about 0.1 Hz.

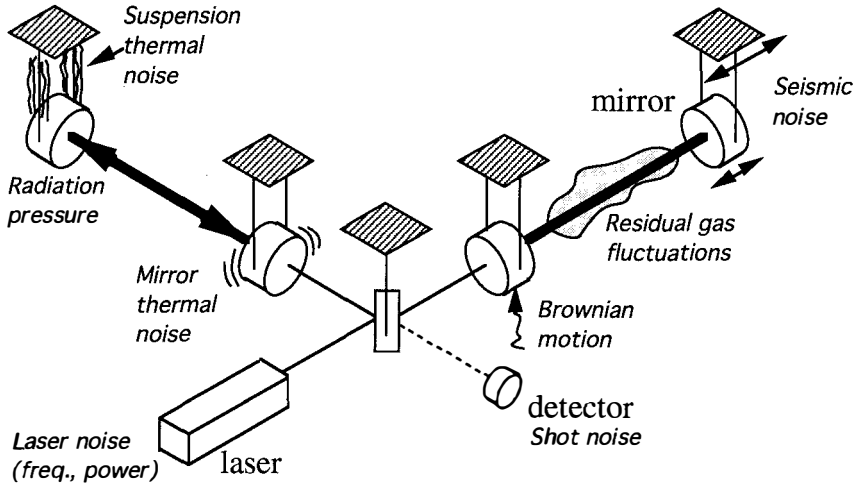


Fig. 2: Overview over some important noise sources.

Thermal noise: A thermal energy of 1/2 kT is stored in each degree of freedom of a thermally equilibrated system, exciting each vibrational mode with a certain amplitude. Especially internal modes of the mirrors causing the whole surface to vibrate back and forth will mimic an arm length change. The most disturbing effect is a wide band noise floor below the resonance itself; it is in fact the sum of the noise floors of all internal mirror modes which limits the VIRGO sensitivity in the medium frequency range[8]. Thermal noise also causes the pendulum, at which the mirror is suspended, to oscillate, and excites violin modes of the suspension wires. Thermal noise can be minimized by selecting materials with a high mechanical quality factor; by selecting a specific mirror shape and weight, the position of the resonance peaks can be manipulated. A heavy mirror (VIRGO: around 20 kg) assures that a given excitation energy leads to a small surface amplitude.

Quantum noise: Phase and amplitude of the light field are conjugate variables in the sense of the Heisenberg uncertainty principle, and cannot be measured simultaneously with arbitrary precision; therefore the power of a light beam falling on a photodiode fluctuates, and the observed photocurrent exhibits a noise (shot noise). The corresponding measurement noise in our case is

$$\tilde{\phi} = \sqrt{\frac{h\,\nu}{P}} \quad \text{or} \quad \tilde{x} = \frac{\lambda}{4\,\pi}\sqrt{\frac{h\,\nu}{P}} \quad ,$$

where $\tilde{\phi}$ and $\tilde{x}$ are the noise spectral densities of light phase and arm length, and P,ν and λ are the laser power, frequency and wavelength. Obviously, the noise can be reduced by increasing the laser power.

A second source of quantum noise comes from the irregular impact of photons on the mirrors, which creates a fluctuating radiation pressure on the mirrors, which in turn perform tiny movements. The noise force F on a mirror of weight M is

$$\tilde{F} = \frac{h\nu}{c}\sqrt{\frac{P}{h\nu}} \quad \text{or} \quad \tilde{x} = \frac{h\nu}{M c \omega^2}\sqrt{\frac{P}{h\nu}} \quad ,$$

where ω is the Fourier frequency. Here, it is best to work at low laser power. The two opposite conditions lead for a given ω to an optimum laser power P_{opt} and a quantum limited displacement noise x_{opt}:

$$P_{opt} = \frac{\lambda M c \omega^2}{4\pi} \; ; \; \tilde{x}_{opt} = \sqrt{\frac{h}{\pi M \omega^2}} \quad ,$$

which gives an $\tilde{x}_{opt}$ of $7 \cdot 10^{-21}$ m/$\sqrt{\text{Hz}}$ (M=10 Kg, $\omega/2\pi$ = 100 Hz); the corresponding laser power of 10 MW is far beyond current technical limits, such that radiation pressure fluctuations play no role in presently planned GW detectors. For VIRGO, it is therefore desirable to have as high a laser power as possible; the foreseen 20 W are chosen such as to be achievable with current technology.

Residual gas pressure fluctuations: Gas in the interferometer arms will lead to apparently varying arm lengths due to refractive index fluctuations. The beam path must therefore be evacuated to a sufficiently low pressure; this is also necessary in order to keep the mirrors clean. The tolerable residual pressure is of the order of 10^{-9} mbar for hydrogen, and 10^{-10} mbar for the other gases. In order to reduce the hydrogen outgassing rate, the steel for the VIRGO tubes will be thermally treated at the factory at 400°C; afterwards, the tube can be baked at 150°C by ohmic heating to remove mostly water vapour. With their kilometer long vacuum tubes for the arms, the GW interferometers presently under construction will represent the biggest existing vacuum systems.

Scattered light noise: Light scattered at mirror surfaces due to imperfections may hit the wall of the vacuum tube, where it picks up seismic noise, since the tube is directly mounted on the ground. This noisy light may be reinjected into the main beam path upon a second scattering process on the other mirror of the arm cavity, which, according to estimates, is not at all negligible. It is therefore important to have extremely high quality mirrors with low scattering losses (0.5 ppm for the VIRGO mirrors); moreover, it will probably be necessary to install baffles in the tubes which trap and absorb scattered light.

Laser noise: In the ideal case of perfect symmetry between both arms, the Michelson interferometer is not sensitive to laser noise, which cancels at the output. Every asymmetry, however, reintroduces a certain degree of sensitivity to laser imperfections. Therefore the laser beam entering the interferometer must be of good quality (low frequency and power fluctuations, and good beam geometry). In order to combine this requirement with the necessary high power, which is difficult with a single laser, the high-quality output of a 1 W monolithic Nd:YAG ring laser[9] (master laser) is sent into the cavity of a high power (20 W) slave laser (injection locking[10]). By this procedure, the slave laser is phase locked to the

laser, such that it reproduces at a higher power level the low noise characteristics of the master laser.

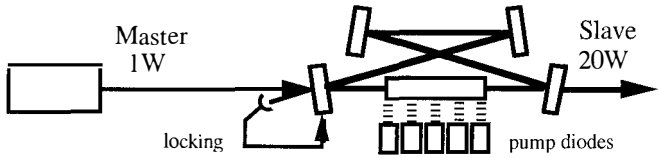


Fig. 3: Generation of a high quality high power laser beam by injection locking.

The required frequency stability of the master laser is achieved by stabilizing to a reference cavity with ultra-low thermal expansion and carefully optimized internal resonances in order to avoid thermal noise. The presently demonstrated frequency stability of 10^{-2} Hz/$\sqrt{\text{Hz}}$ at 500 Hz fulfills the VIRGO specifications. This prestabilized laser will then be further stabilized to the whole of the interferometer as an even more precise frequency reference.

Sending the beam through a special mode cleaner cavity eliminates beam jitter and assures a well defined beam geometry. The lasers operate in the infrared (1.06 μm), because at this wavelength diode-pumped Nd-YAG lasers with good quality and reliability and high power are available.

Other noise sources: Especially disturbing are nonstationary, impulse-like noise signals, since they cannot be distinguished from gravitational waves. Examples are pressure bursts in the vacuum system, electromagnetic disturbances, or sudden release of stored tensions, crystal defects etc. in the mirror suspension wire (creep). For this reason, coincidence measurements between several GW wave detectors are very important; moreover, they give additional information on a GW, such as direction of arrival and polarization.

Overall sensitivity: The expected overall sensitivity is limited by seismic noise below 3 Hz, by pendulum thermal noise in the range up to about 30 Hz, by mirror thermal noise up to 400 Hz, and by shot noise above. The optimum sensitivity will be achieved around 400 Hz with $\tilde{h} \approx 3 \cdot 10^{-23}/\sqrt{\text{Hz}}$ (Fig. 4).

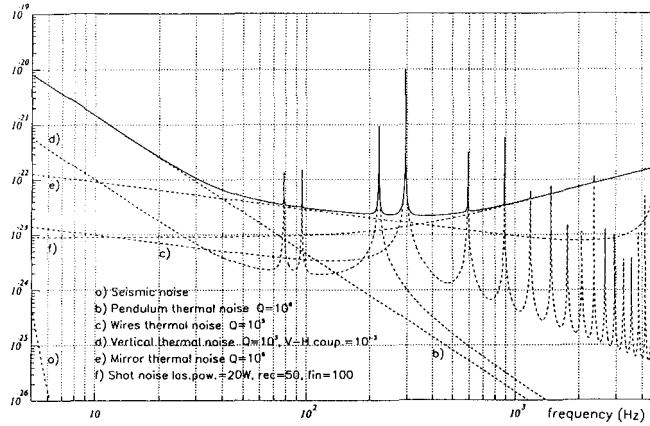


Fig. 4: Sensitivity of VIRGO ($1/\sqrt{\text{Hz}}$), with seismic (a), thermal (b-e) and shot noise (f).

Outlook

The VIRGO interferometer will be constructed near the village of Cascina in Tuscany, not far from Pisa. At present, the site acquisition procedure is in its final phase and the first steps for the construction of buildings are being undertaken. The central area, including laser, mode cleaner, beam splitter and Fabry-Perot near mirrors will be realized first, constituting a small Michelson interferometer of 12 m arm length (without the 3 km arm cavities). With a few changes in the optics, this makes it possible to test a large part of the functioning of the interferometer, such as the control systems, automatic alignment, locking etc.; all this can be done while the 3 km arms are being installed, which will take about 2 years. This strategy makes it possible to shorten the final test phase and to achieve reliable functioning of the interferometer earlier.

References

1) "VIRGO: Proposal for the construction of a large interferometric detector of gravitational waves" (1989).
2) "VIRGO: Final conceptual design" (1992).
3) "VIRGO: Final design" (1995).
4) J. Weber, "Evidence for discovery of gravitational radiation", Phys. Rev. Lett. **22**, 1320 (1969).
5) D. Herriot, H. Kogelnik, R. Kompfner, "Off-axis path in spherical mirror interferometers", Appl. Opt. **3** 523 (1964).
6) J.-Y. Vinet, B. Meers, C.N. Man, A. Brillet, "Optimization of long-baseline optical interferometers for gravitational-wave detection", Phys. Rev. D **38** 433 (1988).
7) R.W. Drever, "Interferometric detectors for gravitational radiation", in Gravitational radiation (Proceedings of the Les Houches Summer Institute), ed. T. Piran, N. Deruelle (1982).
8) F. Bondu, J.-Y. Vinet, "Mirror thermal noise in interferometric gravitational-wave detectors", Phys. Lett. A **198** 74 (1995).
9) T.J. Kane, R.L. Byer, "Monolithic, unidirectional single-mode Nd:YAG ring laser", Opt. Lett. **10** 65 (1985).
10) O. Crégut, C.N. Man, D. Shoemaker, A. Brillet, A. Menhert, P. Peuser, N.P. Schmitt, P. Zeller, K. Wallmeroth, "18W single frequency operation of an injection locked CW Nd:YAG laser", Phys. Lett. **140** 294 (1989).

Résumé

Le détecteur Franco-Italien d'ondes gravitationnelles VIRGO est à présent en cours de construction. Il s'agit d'un interféromètre laser de type Michelson avec des bras de 3 km de longueur, contenant des cavités Fabry-Perot pour augmenter la longueur effective des bras, et utilisant la technique de recyclage de lumière pour augmenter la puissance effective du laser. Ce papier donne brièvement un résumé des différentes sources de bruit limitant la sensibilité des détecteurs interféromètriques d'ondes gravitationnelles, et décrit les principaux éléments constituant l'interféromètre VIRGO.

LIGO PROJECT: AN OVERVIEW

Rencontres de Moriond XXXI
Dark Matter and Cosmology, Quantum Measurements,
Experimental Gravitation

Albert Lazzarini
California Institute of Technology
M/S 51-33 LIGO Project
Pasadena, CA 91125 USA
LIGO-P960007-00-E

ABSTRACT

The Laser Interferometric Gravitational Wave Observatory (LIGO)[1),2)] will open the field of gravitational wave astrophysics. LIGO will detect minute time dependent strains in the spacetime metric induced by gravitational waves using laser interferometry between nearly free masses. The project is designed to permit phased incorporation, at later dates, of improvements in the technology of laser gravitational wave detection to further improve the instrumental sensitivity and bandwidth. LIGO will be part of an international network of long baseline interferometric detectors[3)] to establish the polarization of the waves and the location of the astrophysical sources.

The observations carried out by LIGO are expected to provide fundamental and new information concerning the gravitational interaction including:

- direct measurement of strong field gravity through the observation of the gravitational waves from black holes. The waves will convey information about:
  - the normal modes of black holes
  - inertial frame dragging by rotating black holes
- the observation of compact stellar systems such as neutron star/neutron star, black hole/black hole and black hole/neutron star binaries thereby providing detailed information of the relativistic equations of motion.
- the direct measurement of the polarization states of gravitational waves in conjunction with other interferometric gravitational wave detectors.
- a direct measurement of the speed of propagation of gravitational waves.

The astrophysical information derived from LIGO observations includes:

- the spatial and mass distribution of neutron star binary systems in the universe.
- the spatial and mass distribution of black holes and black hole binary systems in the universe.
- a new and independent method of determining the Hubble expansion using compact binary systems as standard objects.

- the equation of state of neutron stars from the gravitational waveforms at the final coalescence of neutron star binaries.
- the internal dynamics of asymmetric supernova explosions.
- limits to or measurements of the gravitational multipole moments of pulsars.
- limits or observations of the gravitational wave background from the earliest epoch of cosmic evolution.
- a new view of the universe with a high probability of uncovering phenomena not observed by electromagnetic astronomy.

LIGO DETECTOR

The LIGO detector will have the capability of making a confident detection of gravitational waves alone. The LIGO detector will also be capable of providing angle-of-arrival information for detected signals. To achieve these design requirements, the initial detector will consist of three independent laser interferometers operating in coincidence or correlation. Interferometers will be built on two scales: two interferometers have arm lengths of 4 km and the third has arm lengths of 2 km. Each is a Michelson interferometer with resonant Fabry-Perot cavities in the arms. One interferometer will be located at sufficient distance from the other two to enable angle-of-arrival determination by temporal correlation of coincident signals. The two LIGO observatory sites are (i) at the DOE Hanford Nuclear Facility in Washington State (Hanford), and (ii) in Livingston Parish, Louisiana (Livingston).

Hanford houses two instruments in the same vacuum envelope: a full length (4 km) and the half length (2 km) interferometer. Livingston comprises of one 4 km interferometer. The separation between sites is 3000 km, or L/c=10 ms in time delay. The operation of a full length and a half length interferometer at one site serves several functions: [i] it improves the rejection of accidental coincidences by imposing a triple coincidence for a valid burst event; [ii] it is a diagnostic for gravitational waves by demanding that displacement signals scale with interferometer length; [iii] it enables a broad-band search for a stochastic gravitational wave background limited by the environmental correlations at a single site.

The interferometer arms at the two sites are oriented for maximum coincidence sensitivity for a single gravitational wave polarization. This is achieved by orienting one arm of each interferometer at the same angle relative to the great circle passing through the two interferometer sites.

The data from the detector will be analyzed for impulsive, chirp, periodic and stochastic background gravitational waves. Accurate and precise absolute timing will be provided by the Global Positioning System (GPS). Precision permits narrow coincidence

gates among interferometers and accuracy permits correlation of LIGO data with other detector systems, such as resonant bar detectors, particle (neutrino) detectors, and electromagnetic (γ-ray, x-ray, visible, infrared, and radio) astronomical observatories. Gravitational waves signals will be correlated and this property is used in making observations. Both sites include environmental monitoring systems used to (i) identify terrestrial perturbations to the interferometers to reduce false alarm rates in burst searches; (ii) to measure the background perturbations that could influence periodic and stochastic gravitational wave measurements, and (iii) as diagnostics for interferometer development.

INTERFEROMETER CONFIGURATION

The initial LIGO interferometer configuration is a power recycled Michelson interferometer with optically resonant Fabry - Perot cavities in the 4km (2km) arms[4), 5), 6)] (see Figure 1). All the optical components in the phase sensitive part of the interferometer are suspended as pendula to reduce the coupling to seismic and thermal noise and to provide a means to control the optical path lengths in the interferometer. The path lengths are maintained by servo systems to hold the light incident on the detector, placed at the antisymmetric port of the beamsplitter, at a dark fringe. A gravitational wave disturbs this condition by inducing an differential path length change in the two Fabry - Perot cavities and increasing the intensity at the photodetector. When the interferometer is operated at a dark fringe, the light not absorbed or scattered into higher cavity modes by optical components is reflected by the interferometer towards the laser from the beamsplitter. A recycling mirror is placed between the laser and the interferometer to enclose the entire interferometer in an optical cavity. The position and reflectivity of this mirror is chosen so that the light from the laser is added constructively with that circulating in the interferometer while the light reflected by the mirror back to the laser is combined destructively with light emerging from the interferometer. The result is that the optical power circulating in the

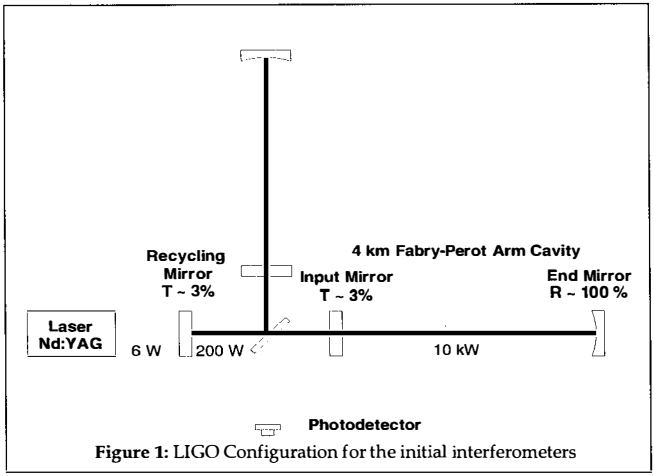


Figure 1: LIGO Configuration for the initial interferometers

interferometer is increased by the reciprocal of the interferometer optical loss. The improvement in shot noise is equivalent to using a more powerful laser. Tables 1 and 2 present the design parameters of the presently planned (initial) interferometers. Also presented in the table are sample parameters for an enhancement contemplated for the initial interferometer. Later phases advanced detector designs will incoporate further configurational changes, such as resonant signal recycling and frequency agile interfeometer designs.

SENSITIVITY

INSTRUMENT NOISE FLOOR

The initial LIGO detector with three interferometers operating in coincidence will have a strain sensitivity adequate to provide confident detection of gravitational waves generated by astrophysical objects and processes discussed above. As an example, the gravitational strain near 100 Hz produced by the coalescence of two 1.4 solar mass neutron stars at a distance of the Virgo cluster has a characteristic magnitude at Earth of $6x10^{-21}$. An initial LIGO detector strain sensitivity goal of 10^{-21} RMS, integrated over a 100 Hz bandwidth centered

Table 1: LIGO Interferometer Optical Parameters

Optical Characteristics	*Nominal Initial Interferometer*	*Sample Enhanced Interferometer*
Arm Length	4000 m	4000 m
Laser Type & Wavelength	Nd:YAG, λ = 1.064 μm	Nd:YAG, λ =1.064 μm
Input Power into Recycling Cavity, P	6W	100W
Contrast Defect, 1-c	3×10^{-3}	3×10^{-3}
Mirror Loss, L_M	1×10^{-4}	1.3×10^{-5}
Power Recycling Gain	30	380
Arm Cavity Storage Time, τ_{Arm}	8.8×10^{-4} s	1.3×10^{-3}s
Cavity Input Mirror Transmission, T	3×10^{-2}	2×10^{-2}
Total Optical Loss, L_T = (Absorption + Scattering)	4×10^{-2}	3×10^{-3}

Table 2: LIGO Interferometer Mechanical Parameters

Mechanical Characteristics	*Nominal Initial Interferometer*	*Sample Enhanced Interferometer*
Mirror Mass, M_M	10.7 kg	40 kg
Mirror Diameter, D_M	0.25 m	0.40 m
Mirror Internal Q_{MP}[7),8)]	1×10^6	3×10^7
Pendulum Q_P(damping mechanism)	1×10^5 (material)	1×10^8 (material)
Pendulum Period, T_P	1 s (Single)	1 s (Double)
Seismic Isolation System[9)]	T(100 Hz) = -100 dB	T(10 Hz) = -100 dB

at the minimum noise region of the strain spectral density is commensurate with this sensitivity. At the threshold of sensitivity, the false signal rate (i.e., triple coincidences not due to astrophysical sources) will be held to less than 0.1 events per year. The design strain spectral densities of the initial instruments and an example enhanced instrument are shown in Figures 2 and 3.

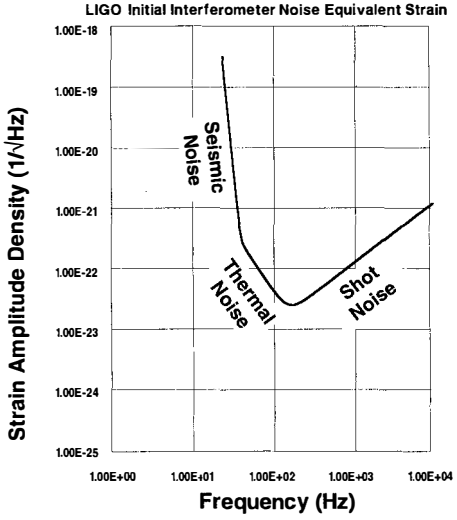


Figure 2: Envelope of the broadband strain spectral density noise floor of the initial LIGO detector. Different spectral regions of the noise floor are dominated by the indicated noise mechanisms.

NATURAL AND FACILITY SENSITIVITY LIMITS

The facilities design permits the ultimately attainable strain sensitivity limits indicated in Figure 4. The limiting sensitivity is set at low frequencies ($f < 20$ Hz) by the naturally occurring gravity gradients due to density variations in the ground and atmosphere. The design goal for the facilities is to not increase the naturally occurring environmental perturbations, such as mechanical vibrations, acoustic noise, electromagnetic fields and gravitational gradients, by more than a factor of two in the gravitational wave detection band

Ultimate LIGO vacuum levels are derived from the need to maintain optical phase noise due to fluctuations in the residual gas column density in the beam tubes and vacuum chambers at a level at or below an equivalent strain noise of 2×10^{-25} $Hz^{-1/2}$. This is expected to be the limiting facility noise source at the highest frequencies ($100 \text{ Hz} < f < 1$ kHz).

The clear aperture of the beam tubes and vacuum chambers is in part determined by the requirement to maintain optical phase noise produced by scattered light to an acceptable level.

SYSTEM EXTENSIBILITY

LIGO is planned to accommodate multiple interferometers at Hanford and Livingston in the advanced phase. These later instruments will have sensitivities representing reasonable extrapolations of the state of the art of interferometric detection over the 30 year lifetime of the facilities. Hence the design and construction of the facilities must permit (i) eventual

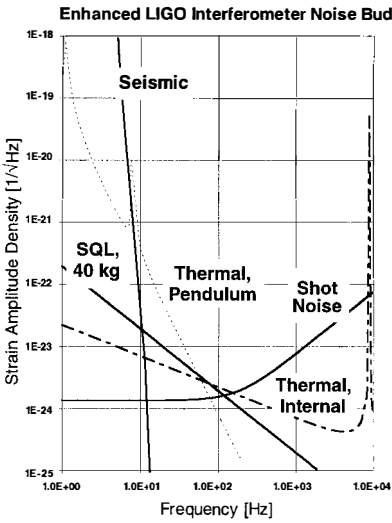


Figure 3: Dominant contributions to the strain spectral density noise floor of an example Enhanced LIGO detector. The Standard Quantum Limit (SQL) shown in Figure 3-3 represents the limiting strain sensitivity arising from radiation pressure fluctuations (producing momentum uncertainty) and shot noise (producing measurement uncertainty). It depends on the mass, sensing frequency, and arm length of the interferometer. The curve shows the locus of points where sensing noise, which varies as $P^{-1/2}$, and radiation pressure, which varies as $P^{1/2}$, are minimized simultaneously].

operation at higher sensitivity and bandwidth and (ii) eventual expansion and upgrade to more interferometers than are in the initial phase.

Beam tube assemblies and vacuum chambers will have clear aperture diameters >1 m capable of accommodating multiple interferometers operating simultaneously.

The layout and design of LIGO facilities constructed in the initial phase will allow further expansion, as required at a later date, to accommodate multiple interferometers with minimum interference during installation and operation. Wherever possible, the construction of the initial phase will provide for extensible infrastructure requiring a minimum of later alteration and replacement other than facility additions.

OBSERVATORY OPERATIONS

Observatory operations must satisfy a range of scientific requirements to further the field of gravitational wave astrophysics. The facility needs to support operations allowing:

- reliable searches for gravitational wave impulsive, chirp, periodic and stochastic background sources at the level of detector sensitivity current at the time;
- retain high enough duty cycle to give a reasonable probability that rare gravitational wave events observable by other gravitational detection techniques or astrophysical measurements will be detected by the LIGO;
- time for the development of improved detectors.

SINGLE INTERFEROMETER OPERATIONS

The goal for the initial detector is the ability to maintain at least one interferometer in operation at an annually integrated availability of 90% with minimum continuous operating periods of 40 hours, allowing for short term loss of lock. Such loss of lock may occur in order to accommodate long-term, low frequency drift (i.e., out of the GW measurement band) by shifting resonant operation from one longitudinal mode to another.

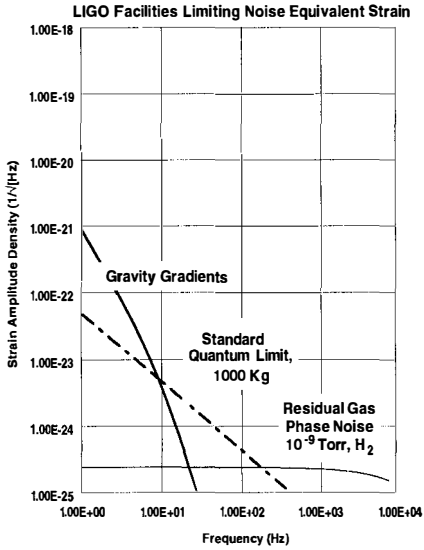


Figure 4: Limiting interferometer performance attributable to the facilities

MULTIPLE COINCIDENCE OPERATIONS

As a goal, LIGO will have the ability to operate in triple coincidence mode for an annually averaged availability of 75%. As a goal, operation in double coincidence mode (defined as operation of the Livingston interferometer with either of the two Hanford interferometers) will be possible with an annually averaged duty cycle of 85%. For both these modes, the minimum period of continuous operation will be 100 hours. The same allowances as above are permitted for short term loss of lock.

DATA

Data collected with the LIGO detector will be available in a format which is compatible with other gravitational wave detectors world wide. This is to enable ease of data interchange and coincidence analysis.

Observatory-based analyses will include immediate-response data for correlation with other (non-interferometric) detectors and also on-line detection of interesting signatures using real-time or near-real-time filtering of the data stream.

LIGO is being constructed jointly by CIT and MIT under NSF Cooperative Agreement No. PHYS-9210038

REFERENCES

1. "Construction, Operation, and Supporting Research and Development of a Laser Interferometer Gravitational-Wave Observatory", Vol. I & II.,LIGO-M890001-00-M. NSF Proposal for LIGO; Technical Supplement to the LIGO Construction Proposal, LIGO-M930002-00-M; LIGO Science Requirements Document, LIGO-E950018.

2. A. Abramovici, *et. al.*, *Science*, Vol. 256, p325.

3. C. Bradaschia, *et. al.*, Nucl. Inst. & Meth. in Phys. Res., A289, 1990.

4. R.W.P. Drever, *et. al.*, Appl. Phys. b31, 1983.

5. D. Shoemaker, *et. al.*, Appl. Optics, Vol. 30, No. 22, 1991.

6. P. Fritschel, *et. al.*, Appl. Optics, Vol. 31, No. 10, 1992.

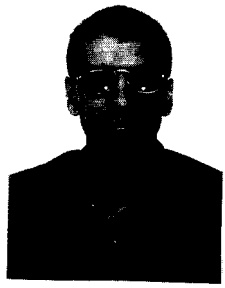
7. P. Saulson, Phys. Rev. D, Vol. 42, No. 8, 1990.

8. A. Gillespie and F. Raab, Phys. Rev. D, Vol. 52, No. 2, 1995.

9. J. Giaime, *et. al.*, preprint, Rev. Sci. Inst.

LASER INTERFEROMETER SPACE ANTENNA (LISA)

Y. Jafry
Space Science Department, ESA/ESTEC, Noordwijk, The Netherlands



Introduction

LISA (Laser Interferometer Space Antenna) is basically a proposal for a "LIGO in space", with the advantages of a longer baseline and a quieter environment than on the ground, allowing the observation of low-frequency gravitational waves (from 10^{-4} to 10^{-2} Hz) which are precluded from terrestrial detectors because of unshieldable gravity-gradient noise associated with seismic motion of matter within the Earth, and with the motion of the atmosphere (better known as 'weather'). The importance of low frequencies is a simple consequence of Newton's laws. For systems involving solar-mass objects, lower frequencies imply larger orbital radii, and the range down to 10^{-4} Hz includes the radii of many galactic neutron star binaries (e.g. the famous Hulse-Taylor binary pulsar) , cataclysmic binaries, some known binaries, and so on. In terms of the likelihood of observation, these are the most certain sources for LISA. Additionally, for highly relativistic systems, where the orbital speeds approach the speed of light, lower frequencies imply larger masses ($M \sim 1/f$), and the range down to 10^{-4} Hz reaches masses of $10^7\, M_\odot$, typical of the black holes that are thought to exist in the centres of many, if not most, galaxies. Their formation and coalescence could be seen anywhere in the Universe and are among the most exciting of possible sources. Detecting them would test the high field limit of gravitational theory and illuminate galaxy formation and quasar models. Signals from compact objects orbiting supermassive black holes and from the postulated gravitational wave cosmological background may also be observable. This rich array of sources is the motivation

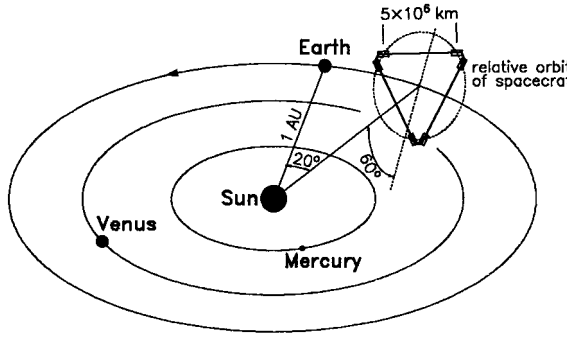


Fig. 1. *LISA interferometer concept: six spacecraft in a triangle, with a pair at each vertex.*

for building a large antenna in space, nicely complementing the high-frequency signals expected from the advanced ground-based detectors.

Mission concept

Conceptual ideas for interferometers in space using separate spacecraft were suggested in the US already in 1978 and 1981. The concept was further developed in the next decade, leading ultimately to the LISA proposal to ESA in 1993. As presently conceived[1)], LISA consists of six identical spacecraft, forming a large equilateral triangle in space (fig. 1). The length of each side of the triangle is 5×10^6 km, defining the interferometer arm length. This is just about the ideal size: if it were larger, the low-frequency gravitational waves would cancel themselves out, resulting in an undesirable loss of science signal; if it were shorter, the spacecraft control requirements would be correspondingly tightened, and thus more difficult to meet. At each vertex of the triangle, there are two closely-spaced (200 km) spacecraft which, together, serve as light sources and beamsplitters for the interferometry. In principle, one "V- shaped" spacecraft at each vertex would be sufficient. However, orbit perturbations will distort the plane of the triangle, such that a single spacecraft at a given vertex would have to incorporate steerable optics (inside each arm of the "V") in order to point its two beams towards each of the two distant vertices at the same time. This complicates the design of each spacecraft, so the alternative "twin spacecraft" approach has been adopted whereby each one uses its attitude control system to point itself (and therefore its single main beam) at the distant vertex. Nevertheless, within each twin pair, a small fraction of the light (10 mW) must be communicated between the neighbouring spacecraft in order for the two lasers to be phase-locked and behave as a single light source. This is achieved using small steerable mirrors. For reasons of compactness, stability, and reliability, LISA will use solid-state diode-pumped monolithic miniature Nd:YAG ring lasers which generate a continuous 1 W infrared beam with a wavelength (λ) of 1.064 μm. Furthermore, with this short wavelength (compared to radio) the light is immune from refraction caused by the charged particles which permeate interplanetary space and thwart the search for gravitational waves by conventional spacecraft radio-tracking. Each main 1 W beam is transmitted to the corresponding remote spacecraft via a 38 cm-aperture f/1 Cassegrain telescope (fig. 2). The same telescope serves to focus the weak return beam (from

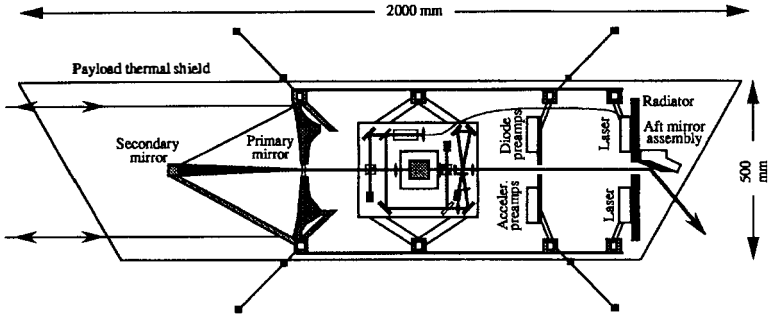


Fig. 2. *Cross-section of the payload on each of the six identical LISA spacecraft. The payload cylinder, surrounded by multiple layers of thermal shielding, houses four major assemblies: a 38 cm diameter f/1 Cassegrain send/receive telescope; an optical bench with laser-injection, cavity for laser stabilisation, beam-shaping optics, photodetectors, and drag-free accelerometer (containing the interferometer "mirror"); a preamplifier disk which carries the accelerometer preamplifiers, photodetector preamplifiers, and an ultrastable oscillator; a radiator disk which carries the lasers (with light fed to the optical bench via a single-mode optical fibre) and radiates their waste heat to space. The small steerable mirror for the laser-link between each corner pair of spacecraft is also mounted on this radiator disk.*

the distant spacecraft) onto a sensitive photodetector where it is superimposed with a fraction of the original local light (and interference takes place). Despite the great distance travelled, the intensity of the received light is high ($\sim 10^8$ photons/sec), making the signal detection straightforward, and many orders of magnitude less demanding than routinely achieved on prototype ground-based interferometers. The interference signals, thus obtained from each arm, are combined in software by the on-board computers to perform the multiple-arm interferometry required to cancel the common-mode phase-noise by nine orders of magnitude. The data are then transmitted to Earth via X-band transponders with steerable 30 cm high-gain antennas. With the triangular configuration, the three arms give two almost-independent interferometers and also provide redundancy in case of the failure of up to two spacecraft (though not at the same vertex).

At the heart of each spacecraft is a vacuum-enclosure containing a polished platinum-gold cube (test mass) which serves as an optical reference ("mirror") for the light beams. When a gravitational wave passes through the system it causes a strain distortion of space, which, in turn, causes fluctuations in the separation of the remotely-spaced test masses. This leads to fluctuations in the optical path between them, causing the phase-shifts which are detected by the interferometry[2]. The goal is to measure distance fluctuations with a precision of 25×10^{-12} m/$\sqrt{\text{Hz}}$ in the frequency range from 10^{-3} and 10^{-1} Hz. Combined with the large separation between the spacecraft, this corresponds to a gravitational-wave strain sensitivity (for stable periodic sources) of 10^{-23} after one year of observation (with a signal-to-noise ratio of five). Obviously, great care must be taken to ensure that the strain measurements are not disturbed by other means. The main error sources fall into two categories: *optical path noise* (e.g. laser shot noise on the receiving photodetectors, laser frequency noise, thermo-mechanical deformation of optical components etc) and *acceleration noise* due to spurious forces acting on

the mirrors. For example, the displacement measurement limit from shot noise is set by the available power, and amounts to about $5\times10^{-12}\,\mathrm{m}/\sqrt{\mathrm{Hz}}$ [1] which is a factor of five below the desired sensitivity. Other optical path disturbances are reduced by the use of ultrastable structures (carbon epoxy with a thermal expansion coefficient of $4\times10^{-7}/\mathrm{K}$) and multilayer thermal insulation on the spacecraft. To eliminate the *acceleration noise*, each test mass is shielded from external disturbances (e.g. solar radiation pressure) by the spacecraft in which it is accommodated. Capacitive sensing is used to monitor the relative motion between each spacecraft and its test mass. These position signals are used in a feedback loop to command thrusters to enable the spacecraft to follow its test mass precisely and without introducing disturbances (in the bandwidth of interest). The ultimate performance of this servo system will be dictated by the noise of the capacitive sensor (e.g. electronic "1/f" noise) and the level of *back-action* and *parasitic* forces (e.g. gravitational, electromagnetic, residual gas pressure etc) acting directly on the test mass. Therefore, the design of the spacecraft and sensor must ensure that these combined effects are below the required level of $10^{-15}\ (\mathrm{m/s^2})/\sqrt{\mathrm{Hz}}$. For historical reasons, this technique of spacecraft control is termed "drag-free control" since it was originally conceived as a means of suppressing atmospheric drag on Earth-orbiting spacecraft[3]. Conventional spacecraft thrusters are not appropriate for the delicate control action required: they are too hefty and too noisy. LISA requires only a few micronewtons to combat the nominal solar pressure, and, furthermore, the thrust must be smoothly controllable to cater for the slight variations in solar radiation flux associated with the internal dynamics of the Sun. Fortunately, suitable thrusters do exist: they are called FEEPs (Field Emission Electric Propulsion), which operate on the principle of reaction from the ejection of ionized caesium atoms[4]. The same thrusters employed by the drag-free system are used to precisely control the attitude of the spacecraft relative to the incoming optical wavefronts, using signals derived from quadrant photodetectors. For a nominal wavefront 'quality' of $\lambda/15$ (non-sphericity), the attitude control must be maintained to within a few $\mathrm{nrad}/\sqrt{\mathrm{Hz}}$, requiring, again, that the sensors (photodetectors) must exceed this noise performance.

Naturally, the LISA spacecraft must be designed to minimise the total mass and required power. Preliminary results yield a mass, per spacecraft, of 300kg, and an operational power requirement, per spacecraft, of 192 W. The six spacecraft, including three propulsion modules for the transfer from the Earth orbit to the final position in interplanetary space, can be launched by a single Ariane 5 rocket. The three pairs of spacecraft are positioned on individual heliocentric orbits of specific inclination and eccentricity in such a way that the three spacecraft-pairs will move relative to each other on a circular orbit inclined at 60° to the ecliptic (fig. 1), maintaining an almost constant separation of the vertices, which is desirable from the point of view of the interferometry (the difference in arm lengths must be known absolutely to within a few hundred metres — easily achievable with X-band radio-tracking). Needless to say, gravitational perturbations from the Earth (and, to a lesser extent, other Solar System bodies) will disturb the orbits. To minimise these effects, the spacecraft constellation should be placed as far away from the Earth as possible. The performance of the Ariane 5 limits the choice to no more than 20° behind the Earth (on its orbit around the Sun), and the resulting orbit perturbations will give rise to undesirable Doppler-shifts in the signals between the various spacecraft. Typical relative motion of about a kilometre per hour can be expected. This

causes a frequency-shift approaching a megahertz, on top of which the phase-measurements are modulated. Compared with the practical data-transmission bandwidth between the spacecraft and the Earth, this frequency shift is much too large, and must somehow be reduced before transmission. Various schemes are under investigation, the most elegant of which is to send two laser-frequencies along each arm (this doesn't require two separate lasers — the second frequency can be modulated as a 'side-band' on the primary). The effects of the Doppler-shifts can be reduced by combining the returning signals such that the phase-information is modulated on the relatively low-frequency obtained from the 'beats' (difference) of the original two Doppler-shifted signals. This can then be further reduced in frequency by 'beating' it against a local ultrastable oscillator to eliminate the remaining shift. An ultrastable oscillator (USO) is a very stable clock which can be used as a frequency reference to mix with the high-frequency Doppler-shifted signal in order to recover the low-frequency science signal. Of course, this clock must be quiet enough that it does not introduce phase-noise during the mixing process. There are already-existing USO's qualified for use on spacecraft. These typically have frequency stabilities of 2×10^{-13} which is good enough if the dual-frequency approach is used. Although, as described, the Doppler-shift due to relative velocities between the spacecraft is undesirable, the velocity of the entire constellation as it orbits the Sun (in the course of one year) can be harnessed and put to good use. The observed gravitational waves are Doppler shifted by the orbital motion. So, for periodic waves with sufficient signal-to-noise, this allows the direction of the source to be determined. It is expected that the strongest LISA sources (from very distant black holes) should be resolvable to better than an arcminute; and even the weaker sources (galactic binaries), to within one degree throughout the entire galaxy.

References

[1] *LISA Pre-Phase A Report* 1996 (unpublished report by the LISA Study Team, available as MPQ Report # 208, Jan 1996)

[2] P.R. Saulson 1994 *Fundamentals of Interferometric Gravitational Wave Detectors* (World Scientific, Singapore)

[3] B.O. Lange 1964 *AIAA Journal* **2(9)** 1590–1606

[4] J. Gonzalez, G. Saccoccia, and H.von Rohden 1993 Paper #IEPC-93-157 (23rd International Electric Propulsion Conference, September, Seattle WA)

TRIAXIAL NEUTRON STARS — A POSSIBLE SOURCE OF GRAVITATIONAL RADIATION

Silvano Bonazzola, Joachim Frieben, and Eric Gourgoulhon
Département d'Astrophysique Relativiste et de Cosmologie, Unité Propre 176 du CNRS,
Observatoire de Paris, F-92195 Meudon Cedex, France

ABSTRACT

Triaxial neutron stars may be important sources of gravitational radiation for the forthcoming generation of interferometric gravitational wave detectors such as LIGO, VIRGO, and GEO600. We investigate the viscosity triggered bar mode secular instability of rapidly rotating neutron stars by means of a perturbation analysis of numerically constructed *"exact"* general relativistic axisymmetric star models. In the theoretical approach, only the dominant parts of the nonaxisymmetric terms of the 3D–Einstein equations are taken into account. A comparison of our results with previous studies of *Newtonian* polytropic stars confirms James' classical result $\gamma_{\rm crit} = 2.238$ for the critical polytropic index. Beyond the Newtonian regime, $\gamma_{\rm crit}$ reveals a slight increase toward highly relativistic configurations. Six out of twelve employed realistic dense matter equations of state admit the spontaneous symmetry breaking for masses above $1.6\,M_\odot$.

1. Introduction

Rapidly rotating neutron stars are highly relativistic objects, and, provided there is some physical process operative, which induces a significant deviation from axisymmetry, may be important sources of gravitational radiation. A transition toward triaxial configurations can

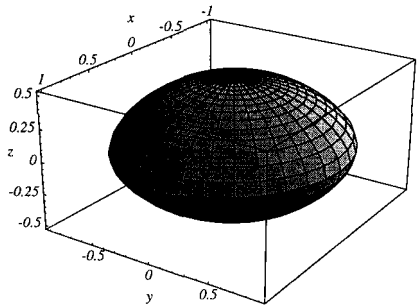


FIG. 1.— Triaxial maximum rotation Newtonian polytropic star for $\gamma = 2.4$. The ellipticity in the equatorial plane is about 0.1. Notice the cusp at the stellar equator in direction of the semimajor axis (x–axis), where the rotation is Keplerian, and which is absent along the semiminor axis (y–axis).

occur, when the ratio $T/|W|$ of rotational kinetic energy and gravitational potential energy reaches some critical value[1,2]. The equation of state, hereafter EOS, has an important influence on the development of a triaxial instability since the neutron star matter must be stiff enough to admit a maximum angular velocity Ω_K higher than the critical one. Homogeneous, incompressible fluid bodies, rotating at moderate constant angular velocity, take the shape of some oblate axisymmetric *Maclaurin spheroid*. Maclaurin spheroids are dynamically unstable for $T/|W| > 0.2738$, but there are two other families of triaxial ellipsoids which bifurcate from the Maclaurin sequence earlier at $T/|W| = 0.1375$. We focus on the sequence of *Jacobi ellipsoids* which rotate uniformly about their smallest axis in an inertial frame. The evolution toward a Jacobi ellipsoid at the bifurcation point is triggered by some *viscous* dissipative mechanism. This instability is called a *secular* one, because it evolves on the associated viscous timescale which is much longer than the dynamical one. The astrophysical realization corresponds to the scenario of a cold highly viscous binary neutron star, being spun up by accretion from its companion. Viscosity dissipates mechanical energy, while it preserves angular momentum. As a consequence, the Maclaurin spheroid, once the instability point is reached, develops toward a Jacobi ellipsoid which exhibits the lowest rotational kinetic energy for a fixed value of the angular momentum. In this final state, viscous dissipation has ceased, and the star is rotating rigidly. Former studies had been performed at the Newtonian[3,4,5,6] or Post–Newtonian[7,8] level which is by no means adequate for highly relativistic objects like neutron stars. Furthermore, the neutron star matter was modeled by a simplifying polytropic equation of state.

2. Numerical models of triaxial neutron stars

Before the symmetry breaking sets in, the neutron star, modeled as a rigidly rotating perfect fluid, can be considered as *stationary* and *axisymmetric*. Making the further assumption that no *meridional* matter currents are present, a favourable choice of *elliptic* field equations was given by Bonazzola et al.[9]. The numerical code[9], based upon these equations, relies on a spectral method, and allows to compute neutron star models with a precision of $\simeq 10^{-14}$ in the spherical symmetric case and $\simeq 10^{-6}$ for maximum rotation configurations, when a $\gamma = 2$ polytropic, *analytic* EOS is used, as well as $\simeq 10^{-4}$ for realistic EOS[10]. The neutron star models are *"exact"* in the sense that the full Einstein equations are solved without any analytic approximation, while the numerical integration covers all space and respects the exact flat space boundary

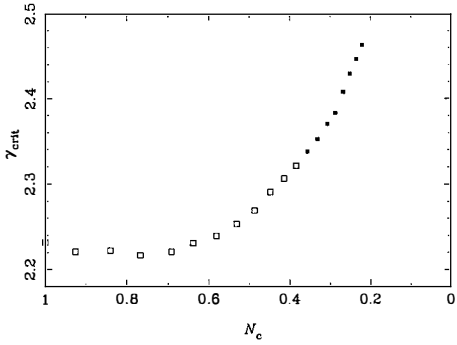


FIG. 2.— Critical polytropic index γ_{crit} as function of the lapse function N_c, measured at the centre of the star. Black boxes indicate configurations unstable with respect to radial oscillations.

condition at spatial infinity. When the symmetry breaking occurs, spacetime is neither stationary nor axisymmetric. However, at the very beginning the deviation from axisymmetry is sufficiently small, and the emitted gravitational radiation may be neglected. Under the additional assumption of *rigid* rotation, the *helical* symmetry of spacetime is preserved. Taking into account only the dominant non–axisymmetric terms, one basically recovers the original equations of the axisymmetric case[11]. The only difference is that the lapse function N — its logarithm $\nu = \log N$ reduces to the Newtonian potential in the weak field approximation — and the matter fields are three–dimensional quantities which depend on r, θ, and $\psi \equiv \phi - \Omega t$, where Ω is the angular velocity of the star. As a consequence, the present approach is exact for the stationary and axisymmetric relativistic, as well as for the fully three–dimensional Newtonian case. After relaxation to a particular axisymmetric configuration, a small perturbation $\delta\nu = \epsilon H_c (r \sin\theta \cos\psi)^2$, which excites the $l = 2$, $m = \pm 2$ bar mode, is added to $\nu = \log N$, and the growth of $\delta\nu$ during the subsequent relaxation is followed. H_c denotes the central log–enthalpy, and ϵ is a small parameter of the order of 10^{-6}. A particular configuration is *secularly unstable* if the perturbation increases, and the fluid body evolves subsequently toward a triaxial spheroid. It is *conditionally secularly stable* if the perturbation tends to zero. In this case, however, three–dimensional terms of higher relativistic order, which have not been taken into account at the present level of approximation, may induce the symmetry breaking.

3. Results for polytropic and realistic equations of state

The investigation of the symmetry breaking of relativistic polytropic stars avoids the problems associated with the use of realistic dense matter EOS[10], and allowed us to perform some comparison with former results obtained for Newtonian stars[4,5,6]. In particular, we have confirmed James'[4] value $\gamma_{crit} = 2.238$ of the critical polytropic index, where secular instability sets in at maximum angular velocity, as well as the asymptotic ratio $T/|W|_{\gamma\to\infty} = 0.1375$. No results in the fully relativistic regime had been reported so far. In the Newtonian case, polytropic stars obey a scaling law, and the critical polytropic index is a global constant. Relativistic effects, however, are supposed to influence the symmetry breaking. Fig. 2 shows the dependence of γ_{crit} on the central value of the lapse function N_c which is an appropriate signature of the relativistic character of the star. After a very slight decrease in the weakly relativistic regime it grows by

EOS	$M_{\max}^{\rm stat}$ $[M_\odot]$	$M_{\max}^{\rm rot}$ $[M_\odot]$	$P_{\rm K}$ [ms]	$P_{\rm break}$ [ms]	$H_{\rm c,break}$	$M_{\rm break}$ $[M_\odot]$
HKP	2.827	3.432	0.737	1.193	0.168	1.886
WFF2	2.187	2.586	0.505	0.764	0.292	1.925
WFF1	2.123	2.528	0.476	0.728	0.270	1.742
WGW	1.967	2.358	0.676	1.042	0.170	1.645
Glend3	1.964	2.308	0.710		stable	
FP	1.960	2.314	0.508	0.630	0.412	2.028
DiazII	1.928	2.256	0.673		stable	
BJI	1.850	2.146	0.589		stable	
WFF3	1.836	2.172	0.550	0.712	0.327	1.919
Glend1	1.803	2.125	0.726		stable	
Glend2	1.777	2.087	0.758		stable	
PandN	1.657	1.928	0.489		stable	

TAB. 1.— Neutron star properties according to various EOS: $M_{\max}^{\rm stat}$ is the maximum mass for static configurations, $M_{\max}^{\rm rot}$ is the maximum mass for rotating stationary configurations, $P_{\rm K}$ is the corresponding Keplerian period, $P_{\rm break}$ is the rotation period below which the symmetry breaking occurs, $H_{\rm c,break}$ is the central log–enthalpy at the bifurcation point and $M_{\rm break}$ is the corresponding gravitational mass. The EOS are ordered by decreasing values of $M_{\max}^{\rm stat}$.

about 10 % for strongly relativistic configurations. One concludes furthermore that *any* neutron star, built upon a polytropic EOS and stable against radial perturbations, becomes secularly unstable for $\gamma > 2.33$ at some critical angular velocity $\Omega_{\rm crit} < \Omega_{\rm K}$. The results for twelve realistic EOS of neutron star matter are presented in Tab. 1. A description of each EOS can be found in Salgado et al.[10], where they had already been employed to construct realistic high precision neutron star models. Six equations turn out to be stable up to the mass–shedding limit $\Omega_{\rm K}$. Only the stiffest ones allow for the symmetry breaking. The correlation with the corresponding maximum mass is not strict because of the density dependence of the EOS. This is notably the case for the WGW EOS which is the stiffest one at moderate densities. For the present sample of realistic EOS, $M_{\rm break}$ has a lower bound of $1.645\,M_\odot$. This is a low value compared

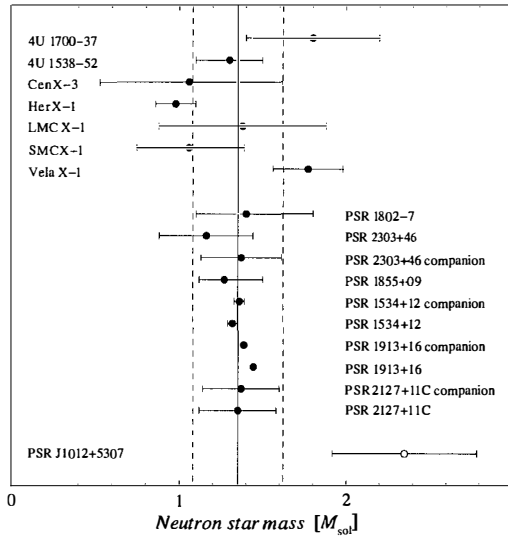


FIG. 3.— Measured masses of 18 neutron stars. Objects in massive X–ray binaries are at the top, radio pulsars and their companions at the bottom. Data compiled by Thorsett et al.[12] is marked by filled circles. The empty circle indicates the mass of the recently discovered radio pulsar PSR J1012+5307, measured by van Kerkwijk et al.[13]. The continuous line indicates the estimate of the mean value of the neutron star mass distribution derived by Thorsett et al. and the dashed lines the corresponding 1σ interval.

with maximum masses of rotating neutron stars built upon a stiff EOS, e.g. in Tab. 1. Unfortunately, the real distribution of neutron star masses is still quite uncertain. Fig. 3 shows the measured masses of 17 neutron stars, compiled by Thorsett et al.[12], and yielding an estimate of the average neutron star mass of $M_{\rm ns} = 1.35 \pm 0.27\,M_\odot$, as well as a recent mass measurement of PSR J1012+5307 of $2.35 \pm 0.434\,M_\odot$[13]. At least the masses of the two X–ray binary neutron stars 4U 1700–37, Vela X–1, and that of the radio pulsar PSR J1012+5307 are well above our minimum mass of $1.645\,M_\odot$. If they prove to be valid after further reduction of the observational errors, we may expect a significant number of neutron stars susceptible of encountering the spontaneous symmetry breaking. Their observation would provide us with useful information about the generation of gravitational waves and constraints on real neutron star matter EOS.

Acknowledgements. J. Frieben gratefully acknowledges financial support by the GOTTLIEB DAIMLER–UND KARL BENZ–STIFTUNG.

REFERENCES

1) Chandrasekhar, S. 1969, Ellipsoidal figures of equilibrium, Yale University Press, New Haven
2) Tassoul, J.L. 1978, Theory of rotating stars, Princeton University Press, Princeton
3) Jeans, J.H. 1928, Astronomy and cosmogony, Cambridge University, Cambridge Press
4) James, R.A. 1964, ApJ, 140, 552
5) Ipser, J.R., Managan, R.A. 1981, ApJ, 250, 362
6) Hachisu, I., Eriguchi, Y. 1982, Prog. Theor. Phys., 68, 206
7) Chandrasekhar, S. 1967, ApJ, 148, 621
8) Tsirulev, A.N., Tsvetkov, V.P. 1982, Soviet Ast., 26, 407
9) Bonazzola, S., Gourgoulhon, E., Salgado, M., & Marck, J.A. 1993, A&A, 278, 421
10) Salgado, M., Bonazzola, S., Gourgoulhon, E., & Haensel, P. 1994, A&A, 291, 155
11) Bonazzola, S., Frieben, J., & Gourgoulhon, E., 1996, ApJ, 460, 379
12) Thorsett, S.E., Arzoumanian, Z., McKinnon, M.M., & Taylor, J.H. 1964, ApJ, 140, 552
13) van Kerkwijk, M.H., Bergeron, P., & Kulkarni, S.R., submitted to ApJ, preprint astro–ph/9606045

ETOILES A NEUTRONS TRIAXIALES — UNE SOURCE EVENTUELLE DE RAYONNEMENT GRAVITATIONNEL

Des étoiles à neutrons triaxiales pourraient être des sources de rayonnement gravitationnel importantes pour la prochaine génération de détecteurs interférométriques d'ondes gravitationnelles comme LIGO, VIRGO ou GEO600. Nous étudions l'instabilité séculaire du mode barre des étoiles à neutrons en rotation rapide par l'intermédiaire d'une analyse perturbative de modèles stellaires axisymétriques *"exacts"* en relativité générale. Dans l'approche théorique on ne tient compte que de la partie dominante des termes non–axisymétriques des équations d'Einstein. Une comparaison avec des études précédentes d'étoiles *newtoniennes* confirme le résultat classique de James à savoir que l'indice polytropique critique vaut $\gamma_{\rm crit} = 2.238$. Au–delà du régime newtonien $\gamma_{\rm crit}$ montre une faible croissance vers les configurations fortement relativistes. Six sur douze équations d'état de la matière dense utilisées admettent la brisure spontanée de symétrie pour des masses supérieures à $1.6\,M_\odot$.

SOME NEW CONCEPTS FOR LASER INTERFEROMETER GRAVITATIONAL WAVE DETECTORS

R.W.P. DREVER
California Institute of Technology, 130-33, Pasadena, California, 91125

ABSTRACT

Some new experimental concepts for extending the performance and range of laser interferometers for gravitational wave detection are described. These include the use of holographic diffraction gratings in test mass mirror coatings to permit higher light power and reduced thermal noise, and possibilities for extending interferometer performance to lower gravitational-wave frequencies by use of magnetic levitation and other techniques.

1. Introduction

The sensitivity of laser interferometer gravitational wave detectors is determined by noise of various types, with the noise source which dominates being dependent on the frequency region being considered. At high gravitational-wave frequencies photon shot noise is generally the limiting factor; at low frequencies seismic noise is expected to set the limit until gravitational gradient noise on the earth is encountered, and at intermediate frequencies thermal noise from the test masses and their suspensions is expected to dominate until quantum limits are reached. We discuss here some experimental concepts intended to facilitate pushing these limits down. To reduce photon shot noise we discuss possibilities for interferometers using diffractive, instead of transmissive, coupling to allow operation at higher light powers, and permit choice of test mass material for lower thermal noise. To reduce seismic noise we introduce some concepts involving magnetic levitation, and also discuss some possibilities for coupling the suspensions at the ends of each arm of an interferometer.

2. Diffractive-coupled interferometers

To minimize photon shot noise high circulating light flux is required in the arms of gravitational-wave interferometers: circulating powers as high as 500 KW have been suggested for "advanced" interferometers using power recycling. Heating effects in even low-loss mirror coatings and in transmission through substrates of beamsplitters or through test masses can be limiting factors. These can be reduced by avoiding transmission through materials, and forming low-amplitude diffraction grating patterns in the mirrors or reflecting coatings. Light is then coupled into optical cavities, or divided and recombined at beamsplitters, by diffraction alone.

2.1 Proposed interferometer configurations

An example of a simple diffraction-coupled Michelson interferometer is illustrated in Fig. 1. Losses associated with blazed gratings may be avoided by choosing a grating spacing which allows only one order of interference, and by using oblique incidence near the Littrow configuration it can be arranged that all diffracted light appears in one output direction, as indicated.

This oblique-angle diffractive coupling technique can be applied to triangular ring cavities, to folded linear Fabry-Perot cavities where the fold occurs at a diffractive mirror, and to complete interferometers. Some possible configurations have been suggested in an earlier paper (1).

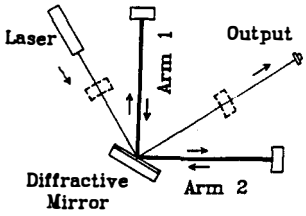


Figure 1

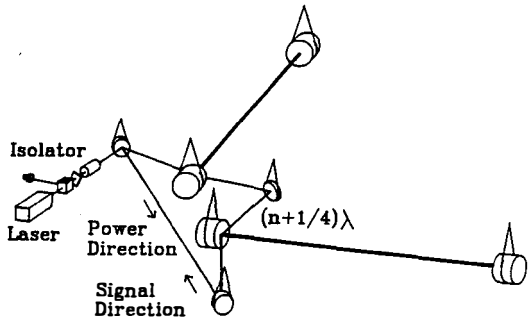


Figure 2

It is possible to use a diffractive mirror at normal incidence, and take advantage of the two symmetrical diffracted beams for input and output coupling. One system of this type is illustrated, in schematic and simplified form, in Fig. 2. Here the two arms of a gravity-wave detector are formed between pairs of suspended non-transmissive test masses, and are coupled via a ring cavity to give some power and signal recycling.

2.2 Advantages of diffractive systems

We originally proposed diffractive coupling for gravity-wave detectors with the main aim of avoiding thermal lensing and allowing increased light power. The possibility of using non-transparent test masses can also make it practicable to use single-crystal test masses of very high mechanical quality factor, Q, thus significantly reducing internal thermal noise. Suitable materials may include silicon and sapphire, which also have the advantage of high thermal conductivity. The removal of substrate absorption losses can also reduce total heating, which may eventually contribute to make it practicable to maintain test masses at reduced temperature, where the highest Q-factors are achieved.

2.3 Considerations on mirror design and manufacture

The low scattering losses desirable make holographic manufacturing techniques advantageous. Some potentially-applicable techniques have been used for other high-power laser systems. However, further development may be required to minimize scattering losses. Techniques may include making a grating profile on the top layer of a mirror coating stack, making the pattern on the substrate, or possibly generating a pattern of refractive index in the coating material by ion implantation or suitable irradiation.

3. Extension of interferometer performance to lower frequencies

Isolation of the test masses of a gravitational wave detector becomes increasingly difficult as operating frequency is reduced, since it is difficult to achieve sufficiently low mechanical resonances in the stages of a passive isolation system, and seismic motions themselves become larger at low frequency. We propose here the possibility of using magnetic levitation to form test mass suspension and isolation systems with lower resonance

frequencies than achieved in simple mechanical systems usually used in this field. We also propose some new coupled isolation concepts.

3.1 Magnetic levitation techniques for test mass suspension

The levitation techniques we are developing are currently all room temperature systems, with the fields generated by permanent magnets. Stabilization of the levitated objects is achieved by sensing vertical position and feeding small control currents to trimming coils around the fixed permanent magnets. Such systems can have negligible power dissipation. We are developing several different magnetic configurations. For test mass suspension it is essential to have very low thermal noise, and correspondingly low dissipative losses. To obtain this, electrically insulating magnetic materials are required, at least in any location where fields depend on the position of the test mass in the direction of the laser beam. Further, coupling to time-varying ambient magnitude fields must be made small enough to avoid introducing significant noise. We propose using one or more pairs of oppositely-oriented magnets on the test mass to cancel dipole and possibly higher-order moments, thus reducing coupling to external fields. An example of a simple 2-magnet configuration we are testing is shown in Fig. 3.

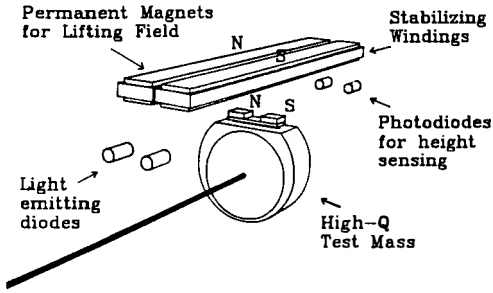


Figure 3

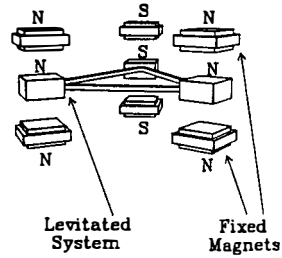


Figure 4

In the version shown, a pair of oppositely-polarized small permanent magnets is attached to a high-Q test mass, via a mechanically-isolating connection to reduce thermal noise. The suspending field is produced by two long, insulating, permanent magnets, with trimming coils. The vertical position of the test mass is sensed by a simple shadow system using two infra-red beams intersecting the top of the test-mass magnets, which control the stabilizing current in the trim coils.

Tests with suspensions of this general type have shown that the period for pendulum-type motion in the direction of the laser beam may be dominated by accidental irregularities in the fields of the suspending magnets. A small trimming permanent magnet (not shown) located to the side of the main magnets may be used to adjust the period. In preliminary experiments, typical periods are around 10 seconds or longer, with relaxation times of several hours.

3.2 Magnetic levitation seismic isolation stages

Slightly different considerations apply here, in that eddy-current damping provided by the conductivity of typical rare-earth magnets can be used to advantage. Further, a system providing more isotropic isolation can be useful. A simple configuration used in some preliminary tests is illustrated in Fig. 4. Here three separate levitated "legs" are used, each having a levitated magnet between two fixed field magnets. Each levitated magnet has its own vertical sensing system, for which we have used Hall-effect sensors (not shown) in this case.

3.3 Tilt-coupled suspensions

The test-mass suspension shown in Fig. 3 is very sensitive to ground tilt when adjusted for a long period. We have earlier suggested coupling the positions of the suspension points of conventional pendulum-type suspensions at each end of the arms of gravitational wave detectors by auxiliary interferometers, to reduce or monitor differential horizontal ground motion. In long-period systems there may be an advantage in coupling tilt motions at the ends of each arm in a similar way, as illustrated schematically in Fig. 4, for the case of a magnetic suspension system.

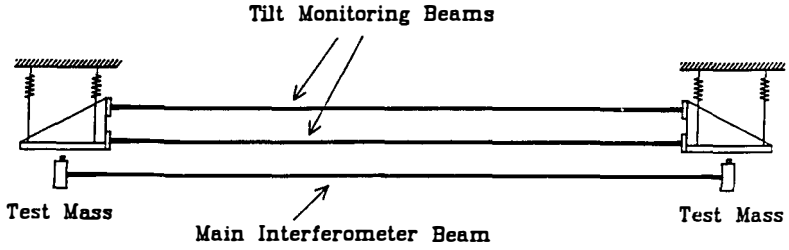


Figure 5

Systems of this general type, whether using magnetic elements or other suspension or isolation techniques seem likely to be useful as a way of improving seismic isolation at very low frequencies.

4. Acknowledgments

I should like to acknowledge the valuable assistance of S. I. Augst, who built and tested several of the magnetic levitation systems investigated. I should also like to thank several colleagues for very helpful discussions, including, in particular, E. W. Cowan, who developed computer models of some of the magnet configurations; and J. L. Hall (JILA) and C. W. Peck for much stimulation and encouragement. The experimental work was supported by the California Institute of Technology.

5. Reference

(1) R.W.P. Drever, in *Proceedings of the Seventh Marcel Grossman Meeting on General Relativity*, Stanford, California, July 1994, eds. M. Keiser and R. T. Jantzen (World Scientific Publishing Co., 1996), p. 1401-1406.

THEORETICAL MOTIVATIONS FOR EQUIVALENCE PRINCIPLE TESTS AND SEARCHES FOR NEW LONG-RANGE FORCES

P. FAYET

Laboratoire de Physique Théorique de l'Ecole Normale Supérieure[1]

[1] Unité Propre du Centre National de la Recherche Scientifique, associée à l'Ecole Normale Supérieure et à l'Université de Paris-Sud.
Postal address: 24 rue Lhomond, 75231 Paris CEDEX 05, France.
e-mail: fayet@physique.ens.fr

Abstract

We discuss the theoretical motivations for new long-range forces induced by spin-1 or spin-0 particles.

The couplings of a new spin-1 gauge particle, the U boson, are obtained by identifying possible extra $U(1)$ gauge symmetries, and taking into account mixing effects with the Z boson. The "fifth force charge" is expressed as $Q_5 = x B + y L + z Q_{el}$ and – within grand-unification – effectively proportional to the number of neutrons. The corresponding force, composition-dependent, would add its effects to those of gravity, leading to apparent violations of the Equivalence Principle, and in general also of the $1/r^2$ Newton's law. Its effective intensity $\tilde{\alpha}$ is related to its range λ and to the extra-$U(1)$ symmetry breaking scale F, by $\tilde{\alpha} \sim 1/(\lambda^2 F^2)$.

An "equivalence theorem" shows that a very light spin-1 U boson does not in general decouple when its gauge coupling vanishes, but behaves like a quasimassless pseudoscalar. Particle physics experiments then constrain F to be larger than the electroweak scale. For large F we get the "invisible U-boson" mechanism.

U-exchanges could also induce spin-spin forces, and even a CP-violating "mass-spin" coupling. These forces, proportional to $1/F^2$, may be more easily detectable than for the axion (for which λ and F are proportionally related).

1 INTRODUCTION.

Why to test the Equivalence Principle? Could new long-range forces exist, in addition to gravitational and electromagnetic ones, and what could be their properties?

The Equivalence Principle is at the basis of the theory of General Relativity. Although we have no reason to believe that general relativity is incorrect, it is certainly not a satisfactory, complete theory. In particular there is a well-known clash between general relativity and quantum physics. More precisely, no consistent quantum theory of gravity exists. While the problem may be ignored temporarily for gravitational interactions of particles at physically accessible energies, it becomes crucial at very high energies of the order of the Planck energy, $\simeq 10^{19}\ GeV$. This is the energy scale at which gravity becomes a strong interaction, so that quantum effects, still ill-defined, become essential. At such energies gravity has an effective intensity comparable to that of the three other interactions, strong, electromagnetic and weak. This is where a unification of all four interactions might occur.

Independently of gravity, the Standard Model of strong, electromagnetic and weak interactions is very successful in describing the physics of elementary particles and their fundamental interactions. But it also suffers from certain difficulties, and leaves a number of questions unanswered. To mention a few:

– it has about 20 arbitrary parameters, including the three gauge couplings of $SU(3)\times SU(2)\times U(1)$, two parameters μ^2 and λ ultimately fixing the W and Higgs boson masses, thirteen mass and mixing-angle parameters associated with the quark and lepton spectrum (plus two others).

– it sheds no light on the origin of the various symmetries and of symmetry breaking, nor on the family problem (why three generations of quarks and leptons, ...).

– in the presence of very large mass scales it suffers from the problem of the stability of the mass hierarchy: how can the W mass remain so small compared to the grand-unification or the Planck scales, in spite of radiative

2 GENERAL FEATURES OF A NEW SPIN-1 INDUCED FORCE.

2.1 Possible extra $U(1)$ gauge symmetries.

For spin-1 particles we can rely on gauge invariance to determine the possible couplings of a spin-1 U-boson, and the expected properties of the corresponding force, should it exist [2]. To do so we first identify the possible extra $U(1)$ symmetries of a Lagrangian density, which are potential candidates for being gauged. This depends crucially on the number of Higgs doublets responsible for the electroweak breaking. In the Standard Model there is no other $U(1)$ symmetry than those associated with baryon and lepton numbers (B and L_i), and with the weak hypercharge Y. More generally, in any renormalizable theory with only one Higgs doublet, any $U(1)$ symmetry generator F must act on quarks and leptons as a linear combination:

$$F = \alpha B + \beta_i L_i + \gamma Y \ . \tag{1}$$

Supersymmetric theories, however, require two Higgs doublets. This leaves room for an additional $U(1)$ invariance, since we may now perform independent phase rotations on these two doublets. With two Higgs doublets separately responsible for up-quark masses, and down-quark and charged-lepton masses, we get:

$$F = \alpha B + \beta_i L_i + \gamma Y + \mu F_{ax} \ , \tag{2}$$

F_{ax} being an extra $U(1)$ generator acting axially on quarks and leptons. Such a combination was gauged, in the first supersymmetric models, to trigger supersymmetry-breaking. These models provided, very early, a natural framework for a possible new long-or-intermediate-range "fifth force" [6, 7], which may now be considered, independently of supersymmetry. Furthermore, in grand-unified theories with large gauge groups including $SU(5)$ or $O(10)$, quarks are related to leptons, so that B and L no longer appear separately, but only through their difference $B - L$. F is then given by:

$$F = \eta \left(\frac{5}{2} (B - L) - Y \right) + \mu F_{ax} \ . \tag{3}$$

corrections which would tend to make it of the same order as m_{GUT} or m_{Planck}?)

– another problem concerns the vacuum energy, when coupled to gravity: if it is not zero, or extremely small, it would generate a too large value of the cosmological constant Λ.

– a delicate question concerns the symmetry or asymmetry between Matter and Antimatter. The CP symmetry (which exchanges matter and antimatter together the left and right orientations in space) is almost a symmetry of all interactions, but it is violated by certain weak interaction effects observed in kaon decays. Once CP is not a symmetry of weak interactions it has no reason to be an exact symmetry of strong interactions. The neutron should then acquire an electric dipole moment, proportional to the parameter θ which measures the effective magnitude of these CP-violating effects of QCD. Since no such moment has been found θ should be smaller than $\approx 10^{-9}$. But why should it be so small if nothing constrains it to vanish? A possible mechanism requires the existence of a new neutral, very light spin-0 particle, the axion [1].

Essentially all attempts to go beyond the Standard Model and try to bring a solution to the above problems involve the introduction of new symmetries, new particles, and new forces. The grand-unification between electroweak and strong interactions would involve very heavy spin-1 gauge bosons that could be responsible for proton decay. The supersymmetry between bosons and fermions requires the existence of new superpartners for all particles. These superpartners – together with the two Higgs doublets required for the electroweak breaking – have a crucial effect on the evolution of the weak, electromagnetic and strong gauge couplings, allowing them to converge, at a large value of the grand-unification scale $\simeq 10^{16}\ GeV$. Supersymmetry is also closely related with gravitation, since a locally supersymmetric theory must be invariant under general coordinate transformations. And superpartners could help solving the hierarchy problem.

More ambitious theories involve extended supersymmetry, new compact space dimensions, and extended objects like superstrings, aiming at a com-

pletely unified description of all interactions. They involve many new particles, including new neutral spin-1 or spin-0 bosons appearing as lower-spin partners or companions of the spin-2 graviton. Spin-1 bosons hereafter called U-bosons could gauge extra-$U(1)$ gauge symmetries, and will be discussed in more detail below [2]. The spin-0 dilaton (or "moduli" ...) fields originating from superstring scenarios may remain massless, and according to a cosmological mechanism attracted towards a point at which they would almost decouple from matter [3]. Their residual interactions could be detected through extremely small (apparent) violations of the Equivalence Principle. This one has already been tested to a very good level of precision ($\approx$ a few 10^{-12}) [4]. The sensitivity could be further improved, down to the 10^{-17} level, by a satellite experiment known as STEP (Satellite Test of the Equivalence Principle) [5].

Should deviations from the Equivalence Principle be detected, further information could allow to distinguish between new spin-1 or spin-0 induced forces, adding their effects to those of gravity. In the first case the new force is generally expected to act on a linear combination of baryon and lepton numbers B and L (i.e. in practice on a combination of the numbers of protons and neutrons, $Z = L$ and $N = B - L$), as we shall see. For spin-0 exchanges the new force is expected to act effectively on a linear combination of B and L with the electrostatic energy. Should such a force be found, testing several pairs of bodies of different compositions could allow to distinguish between the two cases.

Finally we shall discuss forces acting on particle spins, which could be due to the exchanges of spin-1 U bosons or of spin-0 axions. Particularly intriguing is the possibility of a CP-violating interaction that would couple "mass" to "spin".

2.2 Expression of the new "charge" Q_5.

To obtain the current to which the new gauge boson U should couple, we must also take into account mixing effects between neutral gauge bosons. The resulting U current involves a linear combination of the extra-$U(1)$ current with the Z weak neutral current $J_Z = J_3 - \sin^2\theta\ J_{em}$. For simple Higgs systems the extra-$U(1)$ generator, and subsequently the U-current, does not depend on the quark generation considered. The new force should then act on quarks in a generation-independent way. There should be no couplings to strangeness, charm or beauty, nor on mass either. Couplings to a linear combination of B and L with the electrical charge Q, as well as couplings involving particle spins, are expected instead [2, 7]. They will originate from the *vector* and *axial* parts in the U current, respectively.

The vector part in the U current is found to be a linear combination of baryonic and leptonic currents with the electromagnetic current, associated with the charge

$$Q_5 = x\,B + y_i\,L_i + z\,Q_{el} \quad , \tag{4}$$

which reduces to

$$Q_5 = x\,(B - L) + z\,Q_{el} \quad , \tag{5}$$

in the framework of grand-unification. For ordinary neutral matter this further simplifies into

$$Q_5 = x\,B + y\,L = x\,(N + Z) + y\,Z \quad , \tag{6}$$

or, in grand-unification, to

$$Q_5 = x\,(B - L) = x\,N \quad . \tag{7}$$

The action of a spin-1-induced fifth force on neutral matter may then be written in an additive way, proportionally to a linear combination of the numbers of protons (or electrons) and neutrons, Z and N. In the framework of grand-unification, B and L only appear through their difference so that the new force is expected to act effectively on neutrons only. When

the U boson is massless or almost massless and the extra $U(1)$ gauge coupling extremely small, one expects (very small) violations of the Equivalence Principle, since the numbers of neutrons and protons in an object are only approximately proportional to its mass.

Should such a force be discovered, its properties may be used to test its origin, and whether it is due to spin-1 or spin-0 particles: a spin-0-induced force has no reason to act precisely on a linear combination of B and L; it could instead act, approximately, on a linear combination of B and L with the electrostatic energy.

2.3 A relation between range and intensity?

The possible intensity and range of such a new force are largely arbitrary. Still when the U-current has an *axial* part, these two quantities may be constrained, the expected intensity getting smaller when the range gets larger.

The range of the new force may be infinite if the U stays massless. This occurs for example with a single Higgs doublet and no other Higgses, $SU(3)\times SU(2)\times U(1)$ × extra-$U(1)$ being broken down to $SU3)\times U(1)_{QED}\times U(1)_{U-boson}$, the massless U being coupled to a linear combination of the conserved B, L and electromagnetic currents.

In general however the U boson acquires a finite mass, which determines the range of the corresponding force:

$$\lambda = \frac{\hbar}{m_U c} \simeq 2 \text{ meters } \frac{10^{-7}\ eV/c^2}{m_U} \quad . \tag{8}$$

Large masses $\gtrsim$ 200 GeV/c^2 – well in the domain of particle physics – would correspond to extremely short ranges $\lambda \lesssim 10^{-16}$ cm. Small masses of $10^{-5}\ eV/c^2$, $10^{-10}\ eV/c^2$ or less, on the other hand, would lead to macroscopic ranges of 2 cm, 2 kilometers, or more. The new force would then superpose its effects to those of gravitation, leading to apparent deviations from both the Equivalence Principle, and the $1/r^2$ law of gravitation.

The intensity of the new force relatively to gravity is a priori largely

arbitrary. It is characterized (at distances larger than its range λ) by

$$\tilde{\alpha} \approx \frac{g''^2/4\pi}{G_{Newton}\, m_{proton}^2} \simeq 10^{37}\, g''^2 \quad , \tag{9}$$

g'' being the extra-$U(1)$ gauge coupling constant, a priori unknown but which may be extremely small (especially if it is related to gravity).

When it does not vanish the U mass is related to the extra-$U(1)$ symmetry-breaking scale F (determined by the appropriate Higgs v.e.v.'s) by

$$m_U \approx g'' F \tag{10}$$

For a given scale F, the relative intensity of the new force behaves like

$$\tilde{\alpha} \sim g''^2 \sim \frac{m_U{}^2}{F^2} \sim \frac{1}{\lambda^2 F^2} \tag{11}$$

More precisely when model-and-definition-dependent coefficients are taken into account we get the typical estimate [7]

$$\tilde{\alpha} \approx \frac{1}{\lambda\,(\text{meter})^2} \left(\frac{250\ GeV}{F} \right)^2 \tag{12}$$

If the extra $U(1)$ is broken at the electroweak scale ($\approx 250\ GeV$), we get, for example:

$$\tilde{\alpha} \approx \begin{cases} 10^{-5} - 10^{-7} & \text{if} \quad \lambda \simeq 1 \text{ km} \\ 10^{-11} - 10^{-13} & \text{if} \quad \lambda \simeq 10^3 \text{ km} \end{cases} , \tag{13}$$

the latter case being within reach of the planned STEP experiment. But this depends crucially on what the extra-$U(1)$ breaking scale F is.

When the U current is purely vectorial F may be small or even zero for a massless U, and the relative intensity of the new force significantly larger than indicated above in equation (13), making it much easier to detect. But in general the $U(1)$-current also has an axial part, and F will then be constrained directly, from particle physics experiments. Quite remarkably, this occurs even if the extra-$U(1)$ gauge coupling constant g'' is *arbitrarily small*, as we shall see now.

3 A CASE OF EQUIVALENCE BETWEEN SPIN-1 AND SPIN-0 INDUCED FORCES.

3.1 A very light spin-1 U boson does *not* decouple for vanishing gauge coupling – but behaves like a spin-0 particle!

One might think, naïvely, that in the limit of vanishing extra $U(1)$ gauge coupling constant $g"$ the effects of the new gauge boson would be arbitrarily small, and may therefore be disregarded. But this is in general wrong, as soon as a second Higgs doublet is present so that the U-current involves a (non-conserved) axial part!

The amplitudes for emitting a very light ultrarelativistic U boson, proportional to $g"$, seem indeed to vanish with $g"$. However in general it is not so, since the polarization vector for a longitudinal U boson, $\epsilon^\mu \simeq k^\mu/m_U$, becomes singular ($m_U \approx g"$ F also vanishes with $g"$). Altogether the amplitudes for emitting, or absorbing, a longitudinal U boson become independent of $g"$ when the U boson is ultrarelativistic. Such a U boson will then behave very much like a spin-0 particle [6].

This expresses that in the high-energy or low-mass limit ($E \gg m_U$), the third (longitudinal) degree of freedom of a massive U-boson continues to behave like the "eaten" Goldstone boson that gave birth to it. For very small $g"$ the spin-1 U-boson simply behaves as this massless spin-0 Goldstone boson. This applies as well to virtual exchanges. The exchanges of the U boson do not disappear in this limit, owing to the non-conserved axial part in the U current (in general present for more than one Higgs doublet). They become equivalent to the exchanges of a massless (pseudoscalar, CP-odd) spin-0 particle a, having effective axionlike couplings to leptons and quarks

$$2^{1/4}\ G_F^{1/2}\ m_{l,q}\ \ (x \text{ or } 1/x)\ \ \left(r \approx \frac{250\ GeV}{F}\right)\ \gamma_5 \qquad (14)$$

x denoting the ratio of the two Higgs doublet vacuum expectation values.

3.2 "Invisible U-boson" and "invisible axion" mechanisms.

If the extra $U(1)$ gauge symmetry is broken at the electroweak scale by these two Higgs doublets only, the parameter $r = 1$ does not appear in the above formula, and the effective couplings of the spin-1 particle are the same as for the "standard" axion. But this one is excluded from the results of ψ and Υ decay experiments. This is also true for a U boson with non-vanishing axial couplings, if the corresponding $U(1)$ symmetry is broken at the electroweak scale.

Fortunately the ratio r in the above formula, if very small, would make the U boson effects in particle physics practically "invisible". The extra $U(1)$ is then broken "at a large scale" F significantly higher than the electroweak scale ($F \gg 250\ GeV$) – using for example an additional Higgs singlet with a very large vacuum expectation value.

Since a spin-1 U-boson, when very light, is produced and interacts as a spin-0 pseudoscalar axionlike particle, this provided us, at the same time (1980), with a way to make the interactions of the axion almost "invisible", at least in particle physics [6]. This can be realized by breaking the corresponding global $U(1)$ Peccei-Quinn symmetry at a very large scale, through a very large singlet vacuum expectation value, the resulting axion being mostly an electroweak singlet. We then obtained simultaneously both the "invisible U-boson" mechanism, and the "invisible axion" mechanism that became popular later.

3.3 Implications of particle physics experiments for the extra-$U(1)$ symmetry-breaking scale F.

Equation (14) (obtained with two Higgs doublets and an axial part in the U-current) implies that the branching ratios for the radiative production of U bosons – or equivalently of axions – in quarkonium decays are proportional

to r^2 (or equivalently to $1/F^2$):

$$\begin{cases} B\,(\,\psi \to \gamma + U/a\,) \simeq 5\ 10^{-5}\ r^2\,x^2 & C_\psi \\ B\,(\,\Upsilon \to \gamma + U/a\,) \simeq 2\ 10^{-4}\ (r^2/x^2) & C_\Upsilon \end{cases} \tag{15}$$

(C_ψ and C_Υ, expected to be larger than $1/2$, take into account QCD radiative and relativistic corrections). The U boson (quasistable or decaying into $\nu\bar{\nu}$) and the axion (decaying into $\gamma\gamma$ outside the apparatus) would both remain undetected. From the experimental limits [8]:

$$\begin{cases} B\,(\,\psi \to \gamma + \text{“}nothing\text{”}\,) < 1.4\ 10^{-5} \\ B\,(\,\Upsilon \to \gamma + \text{“}nothing\text{”}\,) < 1.5\ 10^{-5} \end{cases} \tag{16}$$

we deduce

$$r \lesssim 1/2 \quad . \tag{17}$$

I.e. the extra-$U(1)$ (local or global) symmetry should be broken at a scale F at least of the order of twice the electroweak scale – using for example extra singlets with large v.e.v's. [6, 7].

For a U with *non-vanishing axial couplings* this can be translated, using equation (12), into an upper bound on the relative strength of the (spin-independent) part of the new force, as a function of its range λ:

$$\tilde{\alpha} \approx \frac{g''^2/4\pi}{G_{Newton}\, m^2_{proton}} \approx \frac{1}{\lambda\,(\text{meter})^2} \left(\frac{250\ GeV}{F}\right)^2 \lesssim \frac{1}{\lambda\,(\text{meter})^2} \quad . \tag{18}$$

This relative intensity is then expected to be rather small, if the range λ is large (e.g. for $\lambda = 10^3$ km, $\tilde{\alpha} \lesssim 10^{-11} - 10^{-13}$, which may still be within the sensitivity of the STEP experiment). On the other hand for shorter λ's (and in particular between 1 mm and more than 10 meters) the relative intensity with respect to gravity is not severely constrained by these arguments, and improved short-range gravity experiments would be worthwhile.

4 NEW FORCES ACTING ON PARTICLE SPINS.

4.1 Spin-spin interactions.

Moreover the exchanges of a new spin-1 U boson would lead to a new, possibly long-ranged, spin-spin interaction potential between two quarks and/or leptons [7, 9]. It is, at distances ρ small compared with the range $\lambda = \hbar/m_U\, c$:

$$\frac{G_F}{8\pi\sqrt{2}}\ \frac{3\ \vec{\sigma}_1 . \hat{\rho}\ \ \vec{\sigma}_2 . \hat{\rho}\ -\ \vec{\sigma}_1 . \vec{\sigma}_2}{\rho^3}\quad (x \text{ or } 1/x)\ (x \text{ or } 1/x)\quad r^2 \qquad (19)$$

($\hat{\rho} = \vec{\rho}/\rho$ denoting the unit vector defined by the two particles), with the constraint $r \lesssim 1/2$ deduced from ψ and Υ decay experiments. The spin-spin potential (19), proportional to $1/F^2$, preserves all P, C and T symmetries and is given by the same expression as for the exchange of a quasimassless axionlike pseudoscalar [10], in agreement with the general equivalence theorem discussed earlier.

This can be understood from the expression of the spin-1 U boson propagator,

$$(g''^2)\ \ \frac{g^{\mu\nu} + \frac{q^\mu q^\nu}{m_U^2}}{q^2 + m_U^2}\ =\ \frac{(g''^2)}{m_U^2}\ \frac{q^\mu\, q^\nu}{q^2 + m_U^2}\ +\ \ldots \qquad (20)$$

For small m_U (or large energies $E \gg m_U$) the dominant contribution arises from the $\frac{q^\mu q^\nu}{m_U^2}$ term, contracted with non-conserved axial contributions to the U current. Moreover $g''^2/m_U^2 \approx 1/F^2 \approx G_F\, r^2$, while $\frac{1}{q^2+m_U^2}$ leads to a Yukawa potential at distance ρ proportional to $\frac{e^{-\frac{\rho}{\lambda}}}{\rho}$. q^μ and q^ν (corresponding to derivative operators in position space) are contracted with axial current contributions proportional to $\bar{q}\,\gamma^\mu\gamma_5\, q$ and $\bar{l}\,\gamma^\mu\gamma_5\, l$ which, in the non-relativistic limit, involve the spin $\vec{\sigma}$ of the quarks or leptons between which U bosons are exchanged. This leads to the spin-spin potential (19).

Both in the spin-1 U-boson and in the spin-0 axion cases, the spin-spin potential is proportional to $1/F^2$. In the case of the axion however (which gets its mass from non-perturbative effects associated with anomalies) there is

a very specific relationship between the range λ and the symmetry breaking scale F:

$$\lambda_{axion} = \frac{\hbar}{m_a\, c} \simeq 2 \text{ mm } \frac{F}{10^{11}\, GeV} \lesssim \text{ cm } . \qquad (21)$$

From astrophysics constraints one usually expects F to be in the $\approx 10^7$ (or 10^{10}) GeV up to $\approx 10^{12}$ GeV range, so that an axion-induced force should in principle be short-ranged ($\lesssim$ cm). For larger values of F and λ the axion-induced force, proportional to $1/F^2$, would be negligible anyway.

On the other hand in the spin-1 U-boson case, the relation between λ and F is different, and involves another parameter, the extra-$U(1)$ gauge coupling $g"$:

$$\lambda_U = \frac{\hbar}{m_U\, c} \approx \frac{1}{g"\, F} \qquad (22)$$

While the axion spin-spin force should be short-ranged, the U-boson spin-spin force may be short or long-ranged. If such a long-range force were found, it cannot be to a spin-0 axion, but to a spin-1 U-boson (or to a spin-0 boson other than an axion, in particular a true massless Goldstone boson).

4.2 A CP-violating "mass-spin" coupling ?

Let us now be even more speculative. If CP-violation effects are introduced, exchanges of a spin-1 U-boson may also lead effectively to a very small CP-violating interaction, resulting in a monopole-dipole force which may be tested in a mass-spin coupling experiment. Again this spin-1-induced force [9] is similar to the one obtained for a quasi-massless spin-0 axion (or axionlike particle) [10].

In the presence of CP-violation effects, we may get – again from the $q^\mu\, q^\nu$ term in the spin-1 propagator – additional contributions originating from the product of a $\vec{\sigma}.\vec{q}$ term for one fermion, times a scalar density for the other. They can mimic an effective interaction resulting from the exchanges of a spin-0 particle having the usual pseudoscalar couplings (14), now supplemented with very small effective (CP-violating) scalar couplings

with quarks. We may use an angle "θ" to parametrize these CP-violation effects, and write the effective CP-violating quark scalar couplings as proportional to $G_F^{1/2}\, r\, m\, \theta$. This leads to an interaction between the spin and the gradient of the Yukawa potential $\frac{e^{-\frac{\rho}{\lambda}}}{\rho}$. The coefficient of this interaction between fermion 1, say an electron of spin $\frac{1}{2}\vec{\sigma}_1$, and a nucleon is proportional to $G_F^{1/2}\, r\, m\, \theta$ (m now denoting an hadronic scale). The resulting CP-violating interaction, at distance ρ, is proportional to

$$G_F\, r^2\, m\, \theta\ \vec{\sigma}_1.\hat{\rho}\ \left(\frac{1}{\lambda\rho} + \frac{1}{\rho^2}\right)\ e^{-\frac{\rho}{\lambda}}\ \approx\ \frac{\theta}{F^2}\ \ldots\ , \tag{23}$$

in which r ($\approx 250\ GeV/F$) is smaller than unity, and θ, presumably $\lesssim 10^{-9}$, measures the magnitude of CP-violating effects. It can be rewritten in terms of effective pseudoscalar and scalar couplings – both proportional to $1/F$ – as

$$\frac{g_p^{(1)}\, g_s^{(2)}}{8\pi\, m_1}\ \vec{\sigma}_1.\hat{\rho}\ \left(\frac{1}{\lambda\rho} + \frac{1}{\rho^2}\right)\ e^{-\frac{\rho}{\lambda}}\ \approx\ \frac{\theta}{F^2}\ \ldots\ . \tag{24}$$

Could such an interaction (which violates the P, CP and T symmetries) be detected in an appropriate experiment searching for a coupling between bulk matter and polarized spins ("mass-spin coupling experiment")? This depends on the range λ of the interaction and of its effective intensity, proportional to $\frac{\theta}{F^2}$ both in the spin-1 and spin-0 cases.

For a spin-0 axion the range λ is proportional to F and expected to be rather short ($\lesssim$ cm, for $F \lesssim 10^{12}\ GeV$). Even in the most favorable case for which $\lambda \sim$ mm, the new force seems too weak to be detected: this would require the CP-violating parameter θ to be $\gtrsim 10^{-5}$, while θ should be $\lesssim 10^{-9}$. And even if λ were allowed to be larger the situation would be worse (the new force being proportional to θ/λ^2). Altogether we cannot yet hope to detect the CP-violating mass-spin coupling induced by axion exchanges.

In contrast the mass of the spin-1 U boson, governed by g" F, may be very small (and λ rather large, $\gtrsim$ cm) even with a symmetry-breaking scale

F not much larger than the electroweak scale. This could make the new "mass-spin" coupling both long-ranged and relatively "intense". The CP-violating spin-dependent potential, still proportional to θ/F^2 (i.e. in this case $\approx g"^2\ \theta\ \lambda^2$) may be significantly larger than for an axion, for which a large symmetry-breaking scale ($F > \frac{1}{2}\ 10^{11}\ GeV$) is required, to have $\lambda \gtrsim 1$ mm.

We now have, with a spin-1 U boson, better chances to detect spin-dependent (dipole-dipole and monopole-dipole) forces. For a spin-0 axion, monopole-dipole CP-violating forces are expected to lie below the level of sensibility of the planned STEP experiment: STEP may test $\theta \gtrsim 10^{-5}$ for a symmetry-breaking scale around $10^{11}\ GeV$ (while θ is expected to be much smaller). For spin-1 particles it may test very small values of θ, significantly lower than the generally-accepted limit of 10^{-9}, provided the extra-$U(1)$ breaking scale is not excessively large compared with the electroweak scale. For a scale $F \approx 10^3\ GeV$ it might test θ down to extremely small values $\approx 10^{-21}$ (for $\lambda \gtrsim 1$ mm). For a scale $\approx 10^8\ GeV$, one could still probe θ down to the $\approx 10^{-11}$ level.

5 CONCLUSIONS.

This leads to a large field of investigations for the search for a new spin-1 or spin-0 induced fifth force, of long or intermediate range. In addition to (apparent) composition-dependent effects of gravitation, and deviations from the inverse-square law, there could be new forces acting on particle spins, which might have detectable macroscopic effects.

References

[1] F. Wilczek, Phys. Rev. Lett. 40 (1978) 279; S. Weinberg, Phys. Rev. Lett. 40 (1978) 223.

[2] P. Fayet, Nucl. Phys. B 347 (1990) 743.

[3] T. Damour and A.M. Polyakov, Nucl. Phys. B 423 (1994) 532.

[4] E.G. Adelberger et al., Phys. Rev. D 42 (1990) 3267; Y. Su et al., Phys. Rev. D 50 (1994) 3614.

[5] STEP, Satellite Test of the Equivalence Principle, Report on the Phase A Study (ESA/NASA, 1993).

[6] P. Fayet, Phys. Lett. 95B (1980) 285; Nucl. Phys. B187 (1981) 184.

[7] P. Fayet, Phys. Lett. B171 (1986) 261; B172 (1986) 363.

[8] C. Edwards et al., Phys. Rev. Lett. 48 (1982) 903;
Crystal Ball Collaboration, Phys. Lett. B251 (1990) 204; CLEO Collaboration, Phys. Rev. D51 (1995) 2053.

[9] P. Fayet, Proc. of the STEP Symposium, Pisa (1993), and to be pub. in Class. Quantum Grav..

[10] J.E. Moody and F. Wilczek, Phys. Rev. D30 (1984) 130.

REDUCING SYSTEMATIC ERRORS IN LLR AND VIKING RADAR RANGING DATA

Kenneth Nordtvedt
118 Sourdough Ridge, Bozeman MT 59715, USA

Lunar laser ranging (LLR) and Viking mission Earth-Mars radar ranging have produced two of the premier solar system data sets available for testing relativistic gravity. But generous allowances for 'systematic errors' in the analysis of this data means that its full potential has not been reached in testing theory. I here briefly sketch work whose general goal is to rationalize the treatment of systematic errors, by including parameterized representations of plausible or suspected biases in either the basic model or a post-model hypothesis. This will allow use of these data sets nearer to the noise limits defined by their model covariance matrices.

During the last 25 years LLR has amassed over 10,000 measurements of the Earth-Moon range. Measurement precision today is typically a couple centimeters and ever-improving, as the active LLR program adds quantity, accuracy and span-of-time to its data set. A key signal sought in the data is a monthly period oscillation which results from any difference in the free-fall rate of Earth and Moon toward the Sun[1] (a so-called Equivalence Principle (EP) violation):

$$\delta r(t) \approx 2.9\cdot 10^{12}\,\frac{a_{moon}-a_{earth}}{g_{sun}}\ centimeters\ \times \left(\cos D + \tfrac{1}{200}\cos 3D + \tfrac{1}{10}\cos(L-D) - \tfrac{1}{20}\cos(D+L) + \ldots\right) \qquad (1)$$

Synodic (D) and anomalistic (L) phases are measured from new moon and from perigee, respectively. Recent analysis of LLR data estimates an amplitude for this signal of -0.7 ± 1.4 cm., the uncertainty being described as "realistic" rather than formal[2]. By comparison, the covariance matrix '2 sigma' uncertainty for this amplitude is only two or three millimeters. But even as it stands, this is our most precise confirmation of universality of free-fall, indicating that Earth and Moon fall toward the Sun at rates equal to 5 parts in 10^{13}. This means, assuming metric gravity, that the Earth's gravitational binding energy, to a part in 10^3 precision, contributes the same to gravitational and inertial mass, confirming the specific non-linear structure of general relativity. But the data has potential to better measure this signal by a factor of 4 or 5 if systematic errors can be brought under more control or modeled.

Oscillations in Earth-Moon range are generally quasi-orthogonal:

$$\int_0^T \frac{\cos \nu t}{\sqrt{T/2}} \frac{\cos \omega t}{\sqrt{T/2}} \, dt \cong \delta_{\nu\omega} + Order\left(\frac{1}{(\nu-\omega)\,T}\right) \tag{2}$$

when sampled uniformly. For technical reasons however, LLR data is sparse near full and new moons, concentrated around quarter phases but skewed toward full moon. Expressing the sum over N observations as a time integral, the data per unit time shows the modulations:

$$\sum_i \rightarrow n_o (1 - C_{2D} \cos 2D - C_D \cos D + \ldots) \, dt \tag{3}$$

with $C_{2D} \approx 0.9$ and $C_D \approx 0.1$. (Data weighting factor modulations may also be relevant). This renders some data signal pairs quite non-orthogonal, a key example being:

$$\sum_{i=1}^{N} \cos D_i \cos 3D_i \cong -\frac{N}{4} C_{2D} \tag{4}$$

The hypothesis --- that there is, <u>by itself</u>, a signal in the data proportional to cosD, the dominant part of the EP-violating signal given in (1) --- therefore has limited meaning and, as I will suggest, should be made part of a more extended hypothesis.

The search procedure for post-model signals in the LLR data begins with formation of a 'residual' vector. The observations are compared with values calculated from a basic model which contains a large number (M > 100) of model parameters:

$$\vec{r} \text{ consists of } r_i = o_i - c(M)_i \quad i = 1 \text{ to } N \tag{5}$$

Small adjustments of the model parameters change the calculated values of the observables in proportion to the parameter 'partials':

$$\frac{\partial \vec{c}}{\partial P_\mu} \equiv \vec{f}_\mu \qquad \text{total model adjustment} = \sum_{\mu=1}^{M} \delta P_\mu \vec{f}_\mu \tag{6}$$

A least squares fit of the model to the observations determines the parameter adjustments and a resulting post-fit (PF) residual vector:

$$\delta P_\mu = \sum_\nu \frac{1}{\vec{f}_\mu \cdot \vec{f}_\nu} \vec{f}_\nu \cdot \vec{r} \quad \text{and} \quad \vec{r}^{\,*} = \vec{r} - \sum_{\mu,\nu} \vec{f}_\mu \frac{1}{\vec{f}_\mu \cdot \vec{f}_\nu} \vec{f}_\nu \cdot \vec{r} \tag{7}$$

Inverse of the model covariance matrix is indicated. The operator which converts the residual vector into a PF residual vector has a useful geometrical interpretation: it is simply the projection operator for the M dimensional vector subspace (M-subspace) defined by the partials vectors of the basic model:

$$\sum_{\mu,\nu=1}^{M} \vec{f}_\mu \frac{1}{\vec{f}_\mu \cdot \vec{f}_\nu} \vec{f}_\nu = \sum_{\kappa=1}^{M} \hat{u}_\kappa \hat{u}_\kappa \equiv \boldsymbol{P_M} \qquad \hat{u}_\kappa \cdot \hat{u}_\lambda = \delta_{\kappa\lambda} \qquad \text{and } \boldsymbol{P_M P_M} = \boldsymbol{P_M} \tag{8}$$

here expressed using an orthonormal representation of the M-subspace:

$$\sum_\mu \delta P_\mu \vec{f}_\mu = \sum_\nu Q_\nu \hat{u}_\nu \qquad \text{with } Q_\nu = \sum_\mu \delta P_\mu \vec{f}_\mu \cdot \hat{u}_\nu \tag{9}$$

If a hypothesis is made that the PF residual vector contains a signal (partials vector $\vec{g}$ plus to-be-fit parameter P') resulting from physics not included in the basic model, the coupled (least squares fit) equations for extra plus model parameters are then:

$$\begin{aligned} P'\vec{g}\cdot\vec{g} + \sum_\mu Q_\mu \hat{u}_\mu \cdot \vec{g} &= \vec{r}\cdot\vec{g} \\ P'\vec{g}\cdot\hat{u}_\nu + Q_\nu &= \vec{r}\cdot\hat{u}_\nu \qquad \text{for each } \nu \end{aligned} \tag{10}$$

To the extent that the extra parameter's partials vector projects into the M-subspace, this parameter is often said to be "correlated" with some basic model parameters, with the further suggestion that biased estimates of those latter basic model parameters will induce bias in the extra parameter. On the face of it, this is suggested by the structure of the coupled equations (10), but final solution for the extra parameter refutes this view, with this parameter estimation occuring <u>outside</u> the M-subspace and therefore being unaffected by whatever biases occur within:

$$P' = \frac{\vec{r}\cdot(1-\boldsymbol{P_M})\cdot\vec{g}}{\vec{g}\cdot(1-\boldsymbol{P_M})\cdot\vec{g}} \equiv \frac{\vec{r}^{\,*}\cdot\vec{g}^{\,*}}{\vec{g}^{\,*}\cdot\vec{g}^{\,*}} \tag{11}$$

Consider a PF residuals vector containing a signal:

$$\vec{I}^* = R_1\vec{g}_1^* + R_2\vec{g}_2^* + \vec{x}^* \qquad (12)$$

(Starred PF vectors have been made orthogonal to the M-subspace by the process in (7)). The first two terms in (12) represent some plausible 'new' physics, while the last term includes all other unknown, unmodeled physics. If one tests a hypothesis consisting of solely the first partials vector of the new physics, one gets:

$$P' = R_1 + \frac{R_2\vec{g}_2^*\cdot\vec{g}_1^* + \vec{x}^*\cdot\vec{g}_1^*}{\vec{g}_1^*\cdot\vec{g}_1^*} \qquad (13)$$

clearly showing the problem with posing a hypothesis which is too limited. Enlarging the hypothesis to include both terms then gives:

$$P' = R_1 + \frac{\vec{x}^*\cdot(\vec{g}_1^* - \vec{g}_1^*\cdot\hat{g}_2^*\hat{g}_2^*)}{\vec{g}_1^*\cdot\vec{g}_1^* - (\vec{g}_1^*\cdot\hat{g}_2^*)^2} \qquad (14)$$

Adding the second signal, the estimate of R_1 becomes immunized to its presence! Such signals should be included in the model or hypothesis. Note also from (14) that changing the data set induces changes in an estimated parameter only via the unmodeled physics.

This is relevant to the search for the EP-violating signal (1). Testing for its presence <u>by itself</u> gives an estimate biased by any signal in the PF residual vector proportional to cos3D. From the sampling modulations (3) and resulting projections (4) one gets:

$$P_D' \cong R_D - \frac{C_{2D}}{2-C_{2D}}R_{3D} \qquad (15)$$

Simply expanding the hypothesis being tested into the form:

$$Hypothesis = P_D' \cos D + P_{3D}' \cos 3D \qquad (16)$$

eliminates this bias, and nothing less than this expanded two-parameter hypothesis should be tested. And there is now the bonus of having estimates of both the cosD and cos3D amplitudes unbiased by each other's presence. It is plausible that some insufficiently modeled physics, perhaps augmented by experimental procedures which are asymmetric with respect to lunar phase, adds a signal to the PF

residuals vector periodic in synodic phase:

$$(\vec{r})_i \approx \sum_{n=1,2..} R_n \cos nD_i + \sum_{n=1,2..} T_n \sin nD_i + \ldots \qquad (17)$$

Finding an appreciable cos3D signal amplitude then suggests that a comparable cosD signal amplitude exists from sources other than an EP violation, whose signal given by (1) has a cos3D 'sideband' with only about 1/200 the amplitude of the principal cosD signal. It may in fact be necessary to further enlarge the PF hypothesis by adding to-be-fit amplitude parameters for more of the Fourier terms. A cos5D signal in residual (17) but not in the hypothesis (16), for example, gives the biased estimate:

$$P_D' \approx R_D - \frac{C_{2D}^2}{4-2C_{2D}-C_{2D}^2} R_{5D} \qquad (18)$$

To estimate both parameters R_D *and* R_{3D} free of bias from all higher odd Fourier terms in (17), one can use the hypothesis:

$$Hypothesis = P_D' \cos D + P_{3D}' \sum_{m=1}^{\infty} f^{m-1} \cos (1+2m) D \qquad (19)$$

with $f = \left(1-\sqrt{1-C_{2D}^2}\right)/C_{2D}$. (In detail, hypotheses (16) and (19) acquire slight modifications when being orthogonalized to the M-subspace --- the 'star' process of (7)). To help find limitations of and make improvements in the basic model, LLR analysts need all available clues concerning the specific signal structure in the PF residuals vector, such as these unbiased amplitude estimates.

The strength of the unknown $\mathbf{x}^*$ signal in (12) and (14) is also worth knowing, for if it is large only a slight lack of orthogonality with partials vectors of estimated parameters will produce significant biases. An algorithm to extract the (non-noise) signal norm from the PF residuals vector is, for equal intervals between observations:

$$Signal\ Norm = \frac{1}{3}\sum_i \left(4r_{i+1}^* r_{i-1}^* - r_{i+2}^* r_{i-2}^*\right) \qquad (20)$$

with straightforward modification for other situations, such as that of unequal intervals between observations. The variance of (20) indicates threshold for measuring a signal norm as $1/\sqrt{N}$ times the

norm of the data's random noise. For LLR this represents a signal with rms size of two or three millimeters. After obtaining hypothesis parameter estimates, the norm of the hypothesis signal can be compared to the PF residuals signal norm found using the algorithm. If the latter significantly exceeds the former, then improvement of the hypothesis is essential in order to capture within itself more, if not most, of the signal norm. These suggested systematic error reducing procedures do generally increase the covariance matrix 'noise' uncertainties in parameter estimation, but as long as the experiment remains systematic error dominated, net improvement in total parameter estimation uncertainty should result.

The Viking mission data set, accumulated during 1976-82 using active radar transponders orbiting Mars or on its surface, consists of over 1000 Earth-Mars radar range measurements of about 7 meters average precision. This range data has been used to measure the gravitational delay of radar propagation to a part in 10^3 precision[3], and to search for cosmological time variation in Newton's coupling parameter, constraining $dG/Gdt \leq 10^{-11}$ per year[4]. It also has the potential to measure the Sun's gravitational to inertial mass ratio to a scientifically interesting level of precision.

The systematic errors in this data set come from an 'identified' source - perturbation of the Mars (and Earth) orbit by the gravity field of the many unmodeled and unknown bodies in the asteroid belt. Though the collective effects of the asteroid belt on the mean motions of Mars and Earth are modeled, the more variable dynamical perturbations have been less considered, while semi-secular effects have remained especially resistant to modeling. Without knowing the specific orbits and masses of the belt's many hundreds of individual bodies, the known regularities of the belt's motions can nevertheless be exploited to construct a correlated error matrix between range observables. The octupole perturbations give in the approximation of circular orbits (quadrupole perturbations give a similar result, but with doubled frequencies):

$$\begin{aligned} \langle o(t)\, o(t+\tau) \rangle \;&=\; A^2 \exp-\left(\tfrac{1}{2}\rho^2\tau^2\right) \;\times \\ &\left(\begin{array}{l} (R(t)\,R(t+\tau) + \xi^2 T(t)\,T(t+\tau))\cos(\omega-\omega_A)\tau \\ +\xi(T(t)\,R(t+\tau) - R(t)\,T(t+\tau))\sin(\omega-\omega_A)\tau \end{array} \right) \end{aligned} \tag{21}$$

ω *and* ω_A are the orbital frequencies of Mars and 'average' asteroid belt, respectively, ρ is an asteroid belt 'decoherence' rate (from the spread of asteroid orbital frequencies), R(t) and T(t) are geometrical observing factors:

$$R(t) = \hat{\rho}_m \cdot \hat{r}_{me} \quad T(t) = \hat{\tau}_m \cdot \hat{r}_{me} \tag{22}$$

Optimal procedure calls for combining this correlated error matrix with the experiment's random measurement noise to obtain the weighting matrix for the least squares fit procedure:

$$Minimize \sum_{i,j} W_{ij}(o_i - c_i)(o_j - c_j) \quad with \quad W_{ij} = \left[\sigma_i^2 \delta_{ij} + \langle o(t_i)\, o(t_j) \rangle\right]^{-1} \tag{23}$$

The structure of the correlations (21) reflects the fact that the planet's orbital perturbation is a quasi-elliptical motion of fixed axis, period $2\pi/(\omega-\omega_A)$, and unpredictable but only slowly changing amplitude and phase. Though not knowing anything about the perturbation at time t, one does know with high confidence that:

$$Given \quad \delta\vec{r}(t)_m \quad Then \quad \delta\vec{r}(t \pm \pi/(\omega-\omega_A))_m \cong -\vec{r}(t)_m \tag{24}$$

viewed in the properly rotating coordinate system. This knowledge enables the construction of a new 'correlated' observable:

$$o(t)' \cong o(t) + \sum_n c(t)_n o(t_n) \tag{25}$$

which is immune from much of the unknown perturbation, achieving about the same as what the optimal weighting matrix can do. The goal is reduction of these particular biases to below a half meter, and then the more difficult problem of bias in semi-secular signals due to eccentricity and inclination of asteroid orbits must be addressed.

Work supported by the Alexander von Humboldt Foundation and by National Aeronautics and Space Administration Contract NASW-4840.

1) K. Nordtvedt. Icarus **114,** 51 (1995); Phys. Rev. **170,** 1186 (1968).
2) J. O. Dickey, et al. Science **265,** 482 (1994).
3) R. D. Reasenberg, et al. Astrophys. J. **234,** L219 (1979).
4) R. W. Hellings. *General Relativity and Gravitation,* edited by B. Bertotti, et al (Reidel, Dordrecht 1984), p 365.

A LOWER BOUND ON NEUTRINO MASS?*

Ephraim Fischbach
Physics Department, Purdue University, West Lafayette, IN 47907-1396 USA

ABSTRACT

The exchange of massless neutrinos between heavy fermions (e.g. e, p, n) gives rise to a long-range 2-body force. It is shown that the analogous many-body force can lead to an unphysically large energy density in white dwarfs and neutron stars. To reduce the energy density to a physically acceptable value, the neutrino must have a *minimum* mass, which is approximately 0.4 eV/c^2. Some recent questions relating to the derivation of this bound are also discussed.

*Work supported in part by the U.S. Department of Energy.

Two recent papers have re-examined the question of whether the forces arising from the exchange of $\nu\bar{\nu}$ pairs can be detected experimentally [1,2]. In the present paper we summarize the main results in Refs. [1,2], and discuss some more recent work aimed at clarifying the lower bound on the neutrino mass derived in Ref.[2].

We begin by considering the self-energy W of a neutron star in the presence of neutrino-exchange forces, which can be evaluated by using the following formula due to Schwinger [3]:

$$W = \frac{i}{2\pi} Tr \left\{ \int_{-\infty}^{\infty} dE \;\; \ell n[1 + \frac{G_F}{\sqrt{2}} a_n N_\mu \gamma_\mu (1+\gamma_5) S_F^{(0)}(E)] \right\}. \tag{1}$$

Here G_F is the Fermi constant, $a_n = -1/2$ is the weak $\nu - n$ coupling constant, N_μ is the vector current of neutrons, and $S_F^{(0)}$ is the free neutrino propagator. As we discuss below, it may be possible to evaluate W for some choices of N_μ without expanding $\ell n[1 + ...]$ in powers of G_F, although for purposes of deriving a bound on the neutrino mass it is useful to carry out such an expansion. The contribution proportional to G_F^k arises from a Feynman diagram in which k neutrons attach to a neutrino loop, as shown in Fig. 4 of Ref. [1]. Since G_F and R are the only dimensional variables in the problem, it follows that in order G_F^k the self-energy of a cluster of k particles spread out through a spherical volume of radius R is proportional to G_F^k/R^{2k+1}. Since one can form $\binom{N}{k}$ k-body clusters from N neutrons, where $\binom{N}{k} = N!/[k!(N-k)!]$ is the binomial coefficient, it follows that the k-body self-energy $W^{(k)}$ is of order

$$W^{(k)} \sim \frac{G_F^k}{R^{2k+1}} \binom{N}{k} \cong \frac{1}{k!} \frac{1}{R} \left(\frac{G_F N}{R^2} \right)^k, \tag{2}$$

where we have written $\binom{N}{k} \cong N^k/k!$ for $k << N$. For a typical neutron star $G_F N/R^2 = O(10^{13})$, and hence it follows from Eq.(2) that higher-order many-body interactions make increasingly *larger* contributions to $W^{(k)}$. Eventually one is led to a paradoxical situation in which $W = \sum_k W^{(k)}$ exceeds the known mass M of the neutron star. If no other mechanism exists to suppress W, then one is led to the conclusion that neutrinos must have a minimum mass m. With $m \neq 0$ the $\nu\bar{\nu}$-exchange force "saturates", just as the strong interaction force does, and for an appropriate value of m the mass-energy of a neutron star arising from neutrino-exchange would be reduced to a physically acceptable value. This minimum mass (for any species ν_e, ν_μ, or ν_τ) is

$$m \gtrsim \frac{2}{3e^3} \frac{G_F}{\sqrt{2}} |a_n| \rho = 0.4 \; eV, \tag{3}$$

where ρ is the number density of neutrons in a neutron star, and $\ell ne = 1$.

The detailed calculations which suggest the possibility of such a bound are presented in Ref. [2]. We briefly review here a number of possible questions that have been raised. a) One

may ask whether perturbation theory is even valid in the presence of effects as large as those found in Ref. [2]. This is addressed in Ref. [2], but an alternative way of viewing the present calculation is to start with $m \neq 0$. For an appropriate value of m, the self-energy W would be acceptably small, and there would be no question concerning the validity of perturbation theory. As one attempts to pass to the $m = 0$ limit, W/M eventually exceeds unity, and this implies that m cannot be smaller than a certain critical value, namely that given by Eq.(3). Using this approach one can sidestep various problems in perturbation theory, and still arrive at the bound in Eq.(3).

b) Another set of issues relates to the possibility of calculating W from Eq.(1) without first carrying out a perturbation expansion. This can be done, for example, in the case of an infinite continuous medium with constant mass density ρ_m. One can anticipate via the following heuristic argument that in such a system the effects of neutrino-exchange will be small. For an infinite system the physically relevant quantity is the energy density which has dimensions μ^4, where μ is a mass scale. The only available dimensional quantities are G_F and ρ_m which appear in the combination $G_F\rho_m \sim \mu^2$. It follows that for an infinite medium we expect to find

$$\text{energy density } \sim (G_F\rho_m)^2. \tag{4}$$

This means that the only contribution in the infinite-medium case is from the 2-body potential, and this conclusion is supported by detailed calculations. By contrast, for a neutron star of radius R with constant number density ρ, one can form the dimensionless quantity appearing in Eq.(2),

$$G_F N/R^2 = (4\pi/3)G_F\rho R. \tag{5}$$

Since the product $G_F\rho R$ is dimensionless it can appear raised to any power in the expression for W, and this is supported by both Eq.(2) and the detailed calculation in Ref. [2]. We can infer from this discussion that large cancellations must take place as one passes from the results for a finite neutron star to those for an infinite medium. Moreover, there are some ambiguities in how the infinite-medium limit is taken, as we show elsewhere. This discussion suggests that calculations of neutrino-exchange effects in an infinite medium may not be directly relevant for a neutron star of finite radius.

c) It has been noted [4] that the same combinatoric factors which enhance the many-body contribution to the self-energy W, also enhance the many-body contribution to the production of physical $\nu\bar{\nu}$ pairs. If these pairs are trapped in the neutron star, then their presence could serve to Pauli-block the exchange of the virtual $\nu\bar{\nu}$ pairs which give rise to the unphysically large value of W. If this were true, then we would no longer be forced to the conclusion that neutrinos must have a minimum mass. However, there are a number of problems with the

preceding scenario: i) For $m = 0$ the analogs of the many-body diagrams considered in Refs. [1,2] would produce both low-energy and high-energy $\nu\bar{\nu}$ pairs at an unphysically large rate. The neutrino star, rather than trapping the neutrinos, would be destroyed as a result of the large forward scattering cross section. ii) Even for low-energy neutrinos or antineutrinos, the dominant many-body interaction may be *repulsive* (depending on the value of N), and hence the neutron star may *expel* both ν and $\bar{\nu}$. (Note that only $k = even$ contributions are non-zero for a spherically symmetric neutron star, and these produce the same effects for ν and $\bar{\nu}$.) More generally, for $m = 0$ problems would arise in both the self-energy W and the $\nu\bar{\nu}$ production rate. To understand how these relate to each other and to other processes occurring in a neutron star (or white dwarf) will require more detailed calculations. However, it seems unlikely that the mechanism proposed in Ref.[4] can avoid the implications of Refs. [1,2] that neutrinos must be massive.

I wish to collectively thank my many colleagues for helpful discussions.

REFERENCES

1. E. Fischbach, D.E. Krause, C. Talmadge, and D. Tadić, Phys. Rev. **D52**, 5417 (1995).

2. E. Fischbach, Ann. Phys. (NY) **247**, xxx (1996).

3. J. Schwinger, Phys. Rev. **94**, 1362 (1954); J.B. Hartle, Phys. Rev. **D1**, 394 (1970).

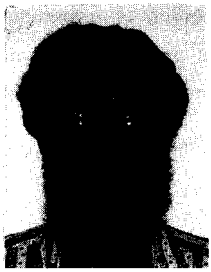
4. A. Yu. Smirnov and F. Vissani, preprint, IC/96/67.

Résumé

L'échange des neutrinos sans masse entre des fermions lourds produit une force macroscopique entre deux corps. Nous montrons que la force analogue entre plusieurs corps mène a une grande densité de l'énergie dons les étoiles des neutrons. Pour réduire la densité de l'énergie à un valeur acceptable, le neutrino doit avoir une masse minimum qui est environ de 0.4 eV/c^2.

PROSPECTS FOR GRAVITATIONAL PHYSICS EXPERIMENTS WITH CRYOGENIC TORSION BALANCES

R.D. Newman, M.K. Bantel and Z.R. Wang
University of California at Irvine
Irvine, California 92717-4575 USA
rdnewman@uci.edu



Abstract

We discuss prospects for improved tests of the weak equivalence principle, searches for anomalies in Newtonian gravity (or new forces), and measurements of the gravitational constant G, using torsion balances operating at cryogenic temperatures. Operation at low temperature promises the benefits of high mechanical Q with correspondingly low thermal noise, high mechanical stability, and high temperature stability. We discuss potential problems in a G measurement due to fiber non-linearity and/or inelasticity, and the problem of minimizing Newtonian gravitational couplings to a balance used for an equivalence principle test. Sensitivity to a differential acceleration of test masses at a level 10^{-15} cm/s^2 or less appears to be a reasonable goal.

1. Introduction

The classic torsion balance is remarkable in its ability to measure extremely weak slowly varying forces, but is nearing its limits when operated at room temperature. The natural next step is to develop torsion balances which operate at cryogenic temperatures. In a previous publication[1] we have discussed limits on such instruments which might be imposed by thermal, seismic, and readout noise sources, concluding that these limitations are consistent with the measurement of differential accelerations of test masses with sensitivity $\Delta a = 10^{-15}$ cm/s^2 or better using an instrument operating at a temperature of about 4K. Such measurements could improve current limits on the universality of free fall in gravitational fields of the sun, earth, and/or local mass sources by three orders of magnitude or more. In the present paper we address other potential limitations on the performance of torsion balances, especially Newtonian gravitational couplings and possible effects of nonlinear and anelastic behavior of torsion fibers.

2. Choice of fiber material

To minimize thermal noise a fiber material with the highest possible torsional oscillation Q is desired. Materials we have considered include: *tungsten*, which is strong and widely used for room temperature torsion balances; *sapphire*, which in bar form exhibits extremely high Q at low temperatures and is available in fiber form with diameter as small as 75 μm (larger than ideal); *aluminum 5056*, which in bar form also exhibits high Q at low temperatures, and is favored for use in gravitational wave bar detectors; and *BeCu*, favored by condensed matter experimentalists for torsion pendulums operated at high frequencies and low temperature. Figure 1 displays Q's measured in our lab, as a function of oscillation amplitude, for pendulums operating at about 4K using 25 cm long fibers of: tungsten (diameter d = 20 μm, period τ_0 = 268 s), hardened aluminum 5056 (d = 50 μm, τ_0 = 105 s), and sapphire (d = 70 μm, τ_0 = 97 s). We are currently testing a BeCu fiber. Fused silica, the fiber material of choice for room temperature instruments, is

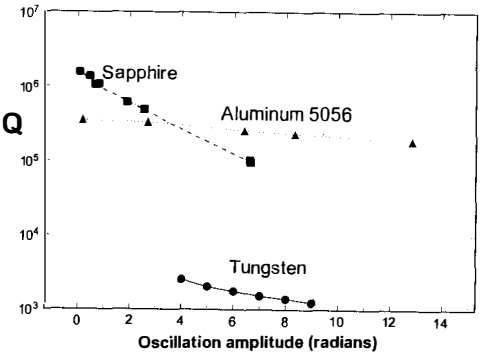


Fig. 1: Measured Q's of torsion pendulums using various fiber materials, as a function of oscillation amplitude at temperatures near 4.2K.

interestingly very unsuitable for use at low temperature, where the temperature coefficient of its shear modulus (and hence its torsion constant) is an order of magnitude greater than at room temperature, while its Q is expected to be little better than at room temperature.

It is surprising that our measured fiber Q's are not higher. For small amplitude oscillation we found a Q of 1,600,000 for sapphire and 360,000 for Al5056, compared with Q's as high as 5×10^9 achieved for sapphire bars at high frequency [2] and 5×10^7 for Al5056 rods at about 1 kHz [3]. Current data on inelastic behavior of materials [4] suggests that torsional Q's would be independent of frequency, (a result of a frequency-dependent effective damping constant $R(\omega) \propto \omega^{-1}$), while comparing our data for aluminum with reference 3 indicates that $Q(1000\ \text{Hz})/Q(.01\ \text{Hz}) \approx 150$. Possible explanations for this include: i) effective damping constants in materials may increase more rapidly with lowered frequency than currently believed, ii) losses at fiber clamp points may dominate our measured Q's, and/or iii) surface effects may dominate loss mechanisms for our thin fibers (it has been found [2] for example that polishing the surface of a sapphire bar greatly improves its Q). It is also possible that the operating temperature of the torsion pendulums we tested was significantly higher than that of the 4.2K helium bath which surrounded the pendulums' vacuum chamber, as a result of heating by the light beam of the optical readout system. We plan further investigation of these questions.

3. Dynamic vs static torsion balance operation modes: Noise sensitivity.

A torsion balance may be used in two ways to probe force fields. When the orientation of the instrument relative to the field sources is changed, one may determine either: I. The resulting static deflection of the balance, either in a fixed or continuously rotating frame, or: II. The resulting change in torsional oscillation frequency of the balance, acting as a torsional pendulum. The first method is that used by Cavendish and by Eötvös, and in most recent applications of the torsion balance to tests of the equivalence principle and searches for new long range forces. The second method has been used in several measurements of G, and by one group [5] in searches for new forces. Each method has distinct advantages; here we consider the relative sensitivity of the two methods to seismic and thermal noise when used to measure a differential acceleration Δa of test masses in a force field.

Let $N^2(\omega)$ be the spectral density of a noise torque acting on the balance or pendulum, and define p_c to be the "composition dipole moment" of the balance (dipole moment of one of the two test mass materials relative to the torsion axis). For a measurement made over a time period t the uncertainty in measured Δa introduced by the noise in the two methods is:

$$\delta a(static) = \frac{\sqrt{N^2(\omega_s)}}{p_c t^{1/2}} \qquad \delta a(dynamic) \approx \frac{\sqrt{N^2(\omega_0)}}{p_c t^{1/2}}$$

Here ω_s is the frequency with which the source field is varied relative to the balance in the static method, and ω_0 is the natural oscillation frequency of the pendulum in the dynamic method. The first equality is easily demonstrated. The expression for the dynamic case is less obvious, but can be shown to be an excellent approximation if three conditions are met: the oscillation amplitude is near an optimum value of 1.84 radians, the measurement extends over several oscillation cycles, and the noise spectrum $N^2(\omega)$ does not have an extreme ω dependence.

Typically a pendulum oscillation period would be on the order of 100 seconds, while the signal frequency in the static method might be a few hours, so that $\omega_0 \approx 100\omega$. Thus whether the static or dynamic method yields lower noise depends on the ω dependence of the noise torque spectrum. Thermal noise arising from the torsion fiber is expected[4] to produce a torque spectrum $N^2(\text{thermal}) \propto \omega^{-1}$, while rotational seismic noise is likely to produce an effective torque spectrum on a torsion balance on the order of $N^2(\omega) \propto \omega^2$. Thus if the dominant noise torque is *thermal*, the dynamic method is likely to have a signal sensitivity *better* by an order of magnitude compared to the static method, while if rotational *seismic* noise dominates, the dynamic method could be *worse* by two orders of magnitude. It is hard to anticipate the level of very low frequency rotational seismic noise to be expected at a quiet field site, so the choice between dynamic and static methods is difficult to make at this point. The limited information available to us on such noise[6] indicates that the dynamic method is favored. Fortunately, an instrument designed for one mode can be readily adapted for operation in the other mode.

4. Effects of fiber non-linearity and anelasticity on measurements of G

Fiber nonlinearity and/or anelasticity are not a significant source of systematic error in null experiments using a torsion pendulum to test the equivalence principle, but are potential error sources in measurements of G such as we plan at UC Irvine[7] using the dynamic method with large amplitude oscillations. Parameterizing nonlinearity in a fiber as a restoring torque related to angular displacement by $N(\theta) = k_1\theta + k_2\theta^2 + k_3\theta^3$, we have determined the ratios k_2/k_1 and k_3/k_1 for a 50 μm diameter 25 cm long aluminum 5056 fiber at 4.2K by analyzing the oscillation frequency $\omega(\theta_0)$ as a function of oscillation amplitude θ_0, and analyzing the harmonic

content of the oscillation time dependence $\theta(t)$ for large amplitude oscillation. We find the nonlinearity to be remarkably small: $k_2/k_1 = (-3.1 \pm 0.4) \times 10^{-7}$ and $k_3/k_1 = (-2.7 \pm 0.7) \times 10^{-7}$. If no correction were made for the effects of such nonlinearity in a G measurement, the error introduced in G can be shown to be less than 1 ppm for oscillation amplitudes less than about 4.5 radians, increasing to 4 ppm for an oscillation amplitude of 9 radians. For a 25 μm fiber such as we plan to use in our actual G measurement, nonlinearity effects should be even smaller.

Real springs are found to behave as if they were a set of individual damped springs in parallel, each with its own characteristic relaxation time (the Maxwell model of anelasticity)[8]. Kuroda[9] has noted that such fiber anelasticity can be a significant source of systematic error in measurements of G using the dynamic method. Kuroda evaluated this error for the case discussed by Quinn et al[10] of an anelastic material characterized by a Maxwell model with a continuum spread of relaxation times and a particular choice of relaxation strength parameter. Kuroda also showed that if the imaginary part of the complex torsion constant $k(\omega) = F/x$ is independent of ω, as appears to be approximately the case for real fiber materials, then the fractional bias in G will be equal to $1/\pi Q$.

The continuum Maxwell model is generally accepted as a good model for anelastic material behavior, but a full determination of the distribution of relaxation strengths for a given material would be impossible. Kuroda's important results raise a concern that for some possible relaxation distribution the anelastic fiber behavior might introduce an undetected G bias significantly bigger than $1/\pi Q$. Fortunately, we have found that Kuroda's analysis can be extended to show that the G bias must be bounded between zero and 1/(2Q) for *any* distribution of relaxation strengths in the Maxwell model. (This upper bound will be correct to order $1/Q^2$). Thus for a pendulum with $Q \approx 360{,}000$, if a correction $\delta G/G = 1/\pi Q$ is made, any residual error associated with anelastic behavior should be less than 1 ppm.

We conclude from these analyses that neither fiber nonlinearities nor anelastic behavior should preclude a measurement of G with accuracy on the order of 1 ppm, using the dynamic method.

5. Pendulum design for a cryogenic equivalence principle (EP) test

The experiment we envision would compare the effective accelerations of two test mass materials, initially probably magnesium and beryllium, in the fields of the earth and sun. The test masses would be two spheres of each material, resting in holes in a beryllium holder (figure 2). Diamond machined mirror faces on the lower end of this holder serve for optical

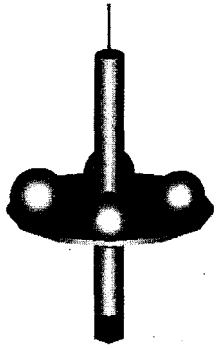


Fig. 2: Pendulum design for a cryogenic equivalence principle test.

position readout. Mass multipole moments of the pendulum are designed to vanish through $\ell=3$. The pendulum would carry 7 gram masses at a radius of 2.4 cm giving a composition dipole moment p_c of 24 g·cm. It would use a 50 μm aluminum 5056 fiber with torsion constant 0.7 dyn·cm/rad, giving a torsional period of about 100 seconds, and would operate in the dynamic mode. As discussed in reference 1, thermal, seismic, and readout noise considerations should allow such an instrument to measure differential accelerations at a level of about 10^{-15} cm/s^2. The key remaining question is how well systematic error from Newtonian gravitational couplings can be controlled. The chief potential source of such error is the coupling of small (nominally zero) m=1 mass multipole moments of the pendulum to gravity field gradients. In principle these moments may be made arbitrarily small by an iterative process in which the pendulum is trimmed by removing small amounts of mass after experimentally determining its response to deliberately augmented gradients. In practice this procedure is limited by the reproducibility of the pendulum's mass distribution at low temperature after a cycle of warming, removal for modification, reinstallation, vacuum bakeout, and recooling. The required reproducability in our design should result from the use of a monolithic mass holder machined as a single piece from beryllium, and spherical test masses which on cooling contract either equally or more than beryllium and thus drop slightly in their holes with reproducibility limited only by deviation from sphericity. Balls of this approximate size can be machined to be spherical within about 1 μin (25 nm), and deviations from spherical shape can be mapped [11] to about 0.1 μin (2.5 nm). Assuming the balls' center of mass lie within 1 μm of their geometrical center, and that the balls are replaced within a milliradian of their original orientation, their center of mass should assume their original position within about 2 nm after a warm/trim/cool cycle. The main effect of ball position shift will be through the coupling of a resulting Q_{xz} type quadrupole moment to an ambient gravity field gradient Φ_{yz}, simulating a differential acceleration of the test masses: $\Delta a = Q_{xz}\Phi_{yz}/p_c$ where p_c is the pendulum's composition dipole moment. For a test mass m a distance ℓ from the axis, displaced vertically by a distance ϵ, Q_{xz} will be $m\ell\epsilon$ while p_c will be $m\ell$ so $\Delta a \approx \epsilon\Phi_{yz}$. In our lab we find we can

readily reduce Φ_{yz} to about 5×10^{-10} s^{-2}; at a field site during dry season it should be possible to maintain Φ_{yz} at a level of 10^{-10} s^{-2}. Assuming $\epsilon \leq 2$ nm leads to an error $\Delta a \approx 4\times10^{-17}$ cm/s^2. Thus we are optimistic that Newtonian gravitational couplings can be controlled at a level that would allow sensitivity at a level $\Delta a = 10^{-15}$ cm/s^2 or better.

6. Comparison with STEP goals

Figure 3 shows limits on a long range force coupling to baryon number that might be obtained with a laboratory instrument with Δa sensitivity of 10^{-15} cm/s^2, compared to present limits and to limits which are the target of space experiments such as STEP. Space EP tests have the advantage of a ~980 cm/s^2 acceleration source, compared to a torsion pendulum's 1.6 cm/s^2 effective field. However the earth-based test we plan has a number of compensating advantages. It can be done at a cost perhaps 1% of that of a space experiment, although probably not more quickly. Checks for systematic errors due to thermal, electromagnetic, and gravitational interactions, by deliberately augmenting the error source, can be made in an operating environment at the design sensitivity level; undesired mass multipole moments may be thus identified and corrected. Charge buildup on test masses from cosmic rays is avoided by use of a conducting torsion fiber, and by operation deep underground if desired. Most importantly, tests for anomalies in the gravitational interaction may be made with high sensitivity in a distance range 10^{-1} to 10^5 m, for which a space experiment is poorly suited.

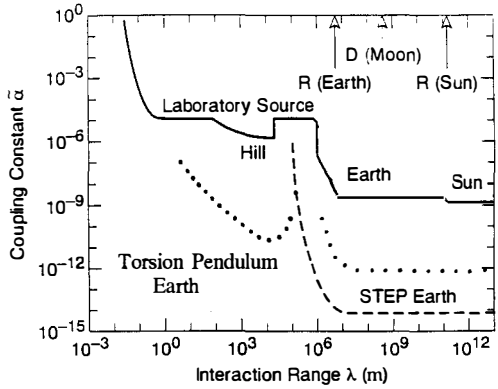


Fig. 3: Limits on a new force coupling to baryon number that might be placed using a torsion pendulum with sensitivity $\Delta a = 10^{-15}$ cm/s^2 (dotted line), compared with existing limits and limits projected for the space-based experiment STEP. Adapted from references 12 and 13.

References

1. R.D. Newman et al, Proceedings of the 7th Marcel Grossmann Mtg, Stanford CA 1994.
2. V.B. Braginsky, V.P. Mitrofanov and V.I. Panov, in *Systems with Small Dissipation*, Univ. of Chicago Press, 1985 p.27,34.
3. William Duffy Jr., J. Appl. Phys. **72**, 5628 (1990).
4. P.R. Saulson, Phys. Rev. **D42**, 2437 (1990).
5. P.E. Boynton et al, Phys. Rev. Lett. **59**, 1385 (1987).
6. Anthony Goodson, Dept. of Physics, U. of Washington, private communication.
7. Michael Bantel et al *Plans for a redetermination of G*, in these proceedings.
8. A.S. Nowick and B.S. Berry, *Anelastic Relaxation in Crystalline Solids*, Academic Press, New York (1972).
9. Kazuaki Kuroda, Phys. Rev. Lett. **75**, 2796 (1995).
10. T.J. Quinn, C.C. Speake and L.M. Brown, Phil. Mag. A**65**, 261 (1992).
11. J.A. Lipa and G.M. Keiser in *Near Zero*, p 641, W.H. Freeman (1988).
12. Y. Jafry, in *Experimental Gravitation*, M. Karim And A. Qadir, eds, p.A301, IOP Publishing (1994).
13. Y. Su et al, Phys. Rev. D**50**, 3614 (1994).

Perspectives Pour Expériences de Physique de Gravitation Avec Balances de Torsion Cryogenique

Résumé

Nous discutons les perspectives pour épreuves perfectionnés du principe équivalence faible, recherches des anomalies de gravité Newtonienne (ou forces nouvelles), et mesurages de la constante de gravitation G, utilisant des balances de torsion fonctionnantes à températures cryogeniques. Fonctionnement à températures basses promet les avantages de haut Q mécanique avec bas bruit thermal, haute stabilité mécanique, et haute stabilité à température. Nous discutons problèmes possibles en mesurant G à cause de non-linearité de la fibre et/ou inélasticité, et le problème à minimiser les accouplements de gravitation Newtonienne à une balance utilisée pour une épreuve du principe d'équivalences. Sensitivité à une accélération différentielle des masses d'épreuve au niveau 10^{-15} cm/s^2 ou moins semble un but raisonnable.

SEARCH FOR ANOMALOUS COUPLING OF SPIN TO HYPOTHETICAL NON-BARYONIC DARK MATTER IN OUR GALAXY

Rogers C. Ritter, Linda I. Winkler*, George T. Gillies
Department of Physics, University of Virginia
Charlottesville, VA 22901, U.S.

**Dept. of Physics and Astronomy, Appalachian State University*
Boone, NC 28608, U.S.

ABSTRACT

A torsion pendulum with special masses having ~10^{23} aligned intrinsic electron spins, but unmeasurable magnetic interactions, was previously used to investigate the existence of weak, hypothetical spin-dependent forces. This instrument has been redesigned to seek a preferential astronomical direction for a long-range spin interaction, such as might result from a nonbaryonic dark matter halo. As the Earth rotates, the pendulum response direction sweeps out the sky, with a patterned observational component along the vector toward the galactic center. After 790 sidereal days of running, a signal lower limit is encountered in conventional analysis, at a level equivalent to about 0.03 of the gravitational acceleration of the masses toward the galaxy center. To study possible systematic contributions to this result, a special analysis uses a "sidereal filter" by which many diurnal artifactual effects should be removed or suppressed. At the present writing, the cumulative mean of the filtered signal towards the galactic center is (0.025 ± 0.011) times the gravitational attraction, where only the statistical error is quoted.

1. INTRODUCTION

Evidence continues to accumulate for the lack of sufficient baryonic matter to account for the flat distribution of stellar velocities in our galaxy. The axion, whose anomalous spin coupling [1] has been studied previously with our pendulum, could be a candidate for the exotic matter to answer this question. But, in view of astrophysical limits on its mass (the Turner window), its range must be too short to couple our Earthbound detector to a significant part of the suggested halo in our

galaxy. Instead, we must look for an unknown spin interaction with longer range.

In related experiments, Phillips [2] and the group of Ni [3] have performed previous torsion pendulum searches for an anisotropy in the interaction of a laboratory spin mass with the cosmic background. These experiments, general searches for a sinusoidal signal in time, and not *a priori* selective in any particular direction, were motivated differently and executed in a less specific way than the present one. In the interest of greater selectivity, the Virginia experiment is a directional search for a match to a pattern predicted from a $\sigma \cdot \mathbf{r}$ interaction with a halo concentric with the center of our galaxy. This is therefore related in some ways to the experiment on composition dependence of the University of Washington group [4].

The mass arrangement in the present experiment is that of a spin quadrupole, while the form of the Phillips and Ni experiments was that of a horizontal dipole. The range dependence arising from the higher multipolarity of the quadrupole, in an interaction for which either multipolarity could couple, would result in an immense sensitivity disadvantage. In view of the null results of the dipole experiments, and the unknown character of any long-range anomalous spin interaction, the present experiment is presented. Here we introduce results of analysis of data encompassing two years of observation.

2. THE EXPERIMENT

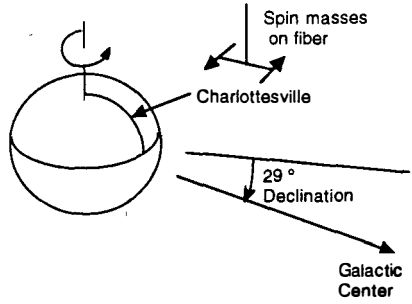


Fig. 1. Mass arrangement on torsion pendulum. Rotation of the Earth scans the sky for regions of stronger anomalous spin interaction with the masses of our detector. A dark matter distribution concentric with the center of our galaxy would torque the pendulum with a 12 - sidereal hour period. Ordinary gravitational sources at such distances, as listed by Stubbs [5], are not seen.

Because of coherence of the huge number of electron spins of these masses in a small independent volume (unlike typical low-temperature spin-aligned targets such

as in accelerator experiments) it is practical to use them in the ultrasensitive macroscopic force sensor, the torsion pendulum. This large spin makes it similar in basic detection sensitivity to NMR measurements used in anomalous spin experiments [6]. The characteristics of our pendulum, and of its DyFe spin masses, have been reported previously [7,8]. In the present experiment, the spin masses are oriented differently, and the pendulum mode was changed from dynamic to static, both to allow the pendulum to sweep out the sky in a directional search efficiently. We continue to use an open geometry, with Helmholz coils to reduce the ambient field at the pendulum below 1 mG, and with temperature measurements for corrections. Ambient magnetic field variations, by comparison with special tests in which the Helmholz coils are ramped, are found to have no measurable effect.

Our effective torque sensitivity is about 10^{-7} dyne-cm for a run period of several days, much like that of the pendulums of Phillips and of Ni. The fiber in the present experiment has a torsion coefficient of 0.09 dyn-cm/radian, about 3 times stiffer than the fibers in those two experiments. Those two tests had a total running time of 13.5 and 10 days, while our computer-operated run covers, thus far, a total of 790 sidereal days (but with occasional unavoidable gaps from power loss, etc.).

3. EXPERIMENTAL RESULTS

The pendulum positional measurement, sampled at 1000 s intervals, and proportional to its torque, provides the basis for the signal of interest. The fit of each 5-day torque sequence to the predicted pattern yields the parameters for the signal, S, for that run. Integration of the point pattern over the halo quoted by Stubbs [5] provides a smoother pattern, but with insignificant consequence in our analysis. Each signal is defined by $S = \varepsilon \cdot R$, in volts, convertible to pendulum torque, and to the corresponding acceleration of the spin masses. Here ε is the standard deviation of the residuals of the fit and R is the dimensionless Pearson correlation fitting coefficient. An individual signal S can be taken to be estimated at an *a priori* one-sigma level.

The mean values, S, of the signal, averaged over accumulated runs, were found to vary from 2 to 4 times the standard deviation of this mean, not decreasing with number of runs N as might be expected from the observed statistical error decrease $\sim N^{-0.5}$. The open geometry and incomplete isolation from ambient effects, particularly temperature variations, suggest local systematic causes as most suspect. To treat these more difficult systematic errors, a new mode of analysis, a "sidereal filter", was devised to separate or suppress diurnal effects, using the 12-sidereal hour

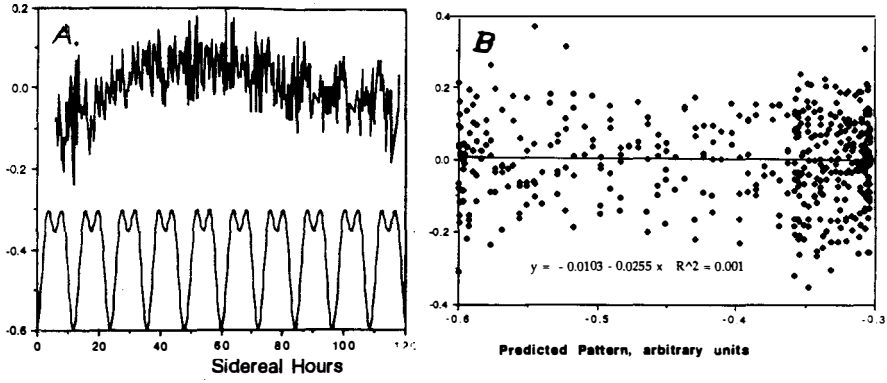


Fig. 2. A. Typical 5-day sequence of data (a standard run), and pattern expected for a $\sigma \cdot \mathbf{r}$ interaction with a halo concentric with galactic center. The scale of the predicted pattern is arbitrary. The signal is in volts of the pendulum position detector (2.7 volts = 0.0175 radian), and has been previously corrected for fiber drift (0.026 μrad/hr), temperature, and time derivative of the temperature.

Fig. 2. B. Five-day torque signal fitted to predicted pattern. The slope of this fit cannot be used to evaluate the results, because of the unscaled pattern.

period of the predicted pattern.

In the new analysis each five-day run is matched with another one starting 183 sidereal days later, which will therefore start one-half day earlier in UTC time. These

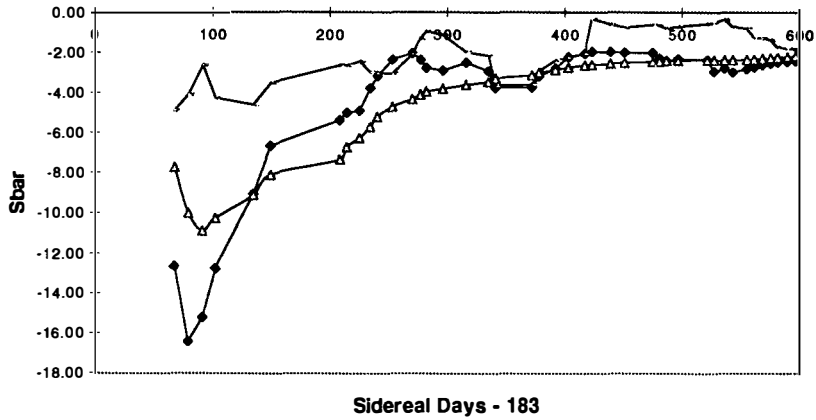


Fig. 3 The accumulation of average values of $\overline{S}^+$ (◆), $\overline{S}^-$ (–), and of $\overline{S}^+ - \overline{S}^-$ (Δ). Data for the first 75 sidereal days are excluded because of major pendulum modifications in this period. Note the trend of $\overline{S}^+$ away from zero.

matched pairs are then summed, S^+, and differenced, S^-, and their comparison constitutes the "filter". Other than doubling, the summing of the data pair should not affect a sidereal signal, but will remove effects that are odd about sidereal t=0, and which might be expected to be a significant part of typical diurnal variations. Conversely, in S^- a true signal would be removed while the odd components of ordinary systematic effects will double. The cumulative sequences of averages of the two analyses, $\overline{S^+}$ and $\overline{S^-}$, respectively, have sequences $\overline{\sigma^+}$ and $\overline{\sigma^-}$ of standard deviations of their means. This paired analysis extracts a statistical cost in the reduction of number of runs, even somewhat more than a factor of two, as a result of our occasional inability to find 183-day matches.

Power spectra, averaged over 30 five-day runs, show peaks at the 1/day and 2/day frequencies, and at some higher frequencies, in both the S^+ and S^- signals. The resolution is far too poor to separate UTC and sidereal peaks, and the high noise level of the S^- spectra (2.5 times that of S^+) hampers accurate sidereal filter evaluation.

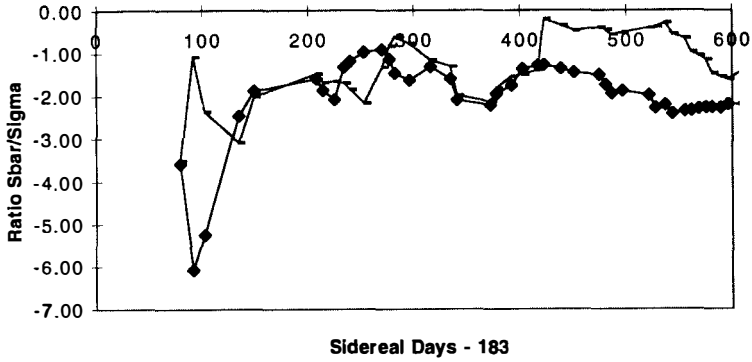


Fig. 4. Trend of accumulated ratios of $\overline{S^+}/\overline{\sigma^+}$ (♦) and $\overline{S^-}/\overline{\sigma^-}$ (–). The $\overline{S^-}/\overline{\sigma^-}$ series has a much rougher profile, and its values are generally smaller than those for the $\overline{S^+}/\overline{\sigma^+}$ series. A constant true signal, or effect which evaded the filtering of our analysis, would show an increasing $\overline{S^+}/\overline{\sigma^+}$ ratio as σ^+ decreases with N.

The series $\overline{S^+}$ shows a pervasive negative torque signal (repulsive in $\sigma \cdot \mathbf{r}$) with a level of 2.3 σ at the present time. This signal level corresponds to a 7×10^{-5} degree rotation of our pendulum, equivalent to an acceleration of either of the spin masses by an amount 4.6×10^{-10} cm/s^2 away from the center of the implied halo. This is 0.025 of the gravitational acceleration of a mass towards the galactic center.

The 2.3 σ level of the experimental signal is an interesting but not satisfactory statement about such an unexpected effect. The present plan is to accumulate approximately 15 more runs. The increasing values of N will smooth the averaged series fluctuations, and thereby establish a little more strongly that there is a trend, if it continues, before a decisive pendulum change. Then we will reverse the sense of rotation of the spin masses about the fiber, which should reverse the sign of any true signal. A sign reversal occurring in the subsequent runs, approximating the present level of uncertainty, 2.3 σ, would correspond to an almost vanishing statistical probability that this occurred by accident. It is hard to predict an artifactual signal with that sidereal behavior, and which would also have such a reversible signature.

REFERENCES

1. Moody, J.E. and Wilczek, F., Phys. Rev. D **30**, 130 (1984).
2. Phillips, P., Phys. Rev. Lett. **59**, 1784 (1987).
3. Wang, S.-L., *et al*, Mod. Phys. Lett. A **8**, 3715 (1993).
4. Smith, G., et al., Phys. Rev. Lett. **70**, 123 (1993).
5. Stubbs, C.W., Phys. Rev. Lett. **70**, 119 (1993).
6. See, e.g. D.J. Wineland *et al*, Phys. Rev. Lett. **67**, 1735 (1992).
7. R.C. Ritter *et al*, Phys. Rev. D **42**, 977 (1990).
8. R.C. Ritter *et al*, Phys. Rev. Lett. **70**, 701 (1993).

RESUME:

Un pendule de torsion avec des masses contenant ~ 10^{23} spins électroniques intrinsèques alignés, mais sans interaction magnétique mesurable, a été utilisé dans le passé pour chercher d'hypothétiques forces faibles dépendant du spin. On a reconfiguré cet instrument pour chercher si une interaction de spin à grande distance existe dans une direction privilégiée, comme celle qui pourrait être due à un halo de matière invisible non baryonique. Grâce à la rotation de la Terre, la direction dans laquelle le pendule est sensible parcourt le ciel et l'angle qu'elle fait avec la direction du centre galactique a une variation périodique. Après 790 jours sidéraux d'observation, l'analyse statistique fixe une limite inférieure pour un tel signal au niveau de 3% de l'accélération gravitationnelle vers le centre galactique. Pour étudier la contribution de possible effets systématiques, on a appliqué un "filtre sidéral" qui doit supprimer beaucoup de ces effets de signature diurne. Au moment où nous écrivons ces lignes, la moyenne cumulative après filtrage d'un signal vers le centre galactique est (0.025 +/- 0.011) fois la valeur de l'attraction gravitationnelle, où l'incertitude rapportée est celle obtenue par le traitement statistique.

PRECISION MEASUREMENTS OF SMALL FORCES

Michael W. Moore and Paul E. Boynton
Department of Physics, Box 351560
University of Washington
Seattle, WA 98195 USA

Abstract. After more than 200 years, the torsion balance/pendulum is still one of the most precise means for measuring weak, macroscopic-range forces. Historically, two primary measurement techniques have been developed for this instrument. In one, the signal is the equilibrium angular displacement of the pendulum resulting from the presence of an external torque. In the other, this torque is detected by the associated fractional change in the natural frequency of torsion oscillations. Both techniques have been refined to the point that some experiments are now limited by temporal variations in torsion fiber temperature, despite special efforts to insure temperature stability.

We have devised and tested another torsion pendulum measurement technique for which the effect of fiber temperature variations is suppressed by more than four orders of magnitude. Our observable is the amplitude of the second harmonic of the pendulum motion. A recent 21-day measurement using a null pendulum (no signal expected) with Q = 4,000 yielded an amplitude of 0.4±0.6 nanoradians for the second harmonic. This is equivalent to an equilibrium pendulum displacement of 1.6±2.4 nanoradians.This new technique suggests that a significant improvement in the precision of these fundamental physics experiments may be realized with a moderately high Q, room-temperature apparatus.

1. Modern Torsion Balance Experiments

For more than 200 hundred years, the torsion balance has been used to make increasingly precise measurements of very weak forces. Recent applications include searches for composition-dependent forces (coupling to baryon number, lepton number, or spin), composition-independent tests of the inverse square law of gravity, and precise measurements of the gravitational constant, G. Various methods have been employed to detect these weak forces. The most common has been a direct measurement of the torsion balance's deflection[1,2], but others have looked at resonantly induced changes in oscillation amplitude[3], and variations in oscillation frequency at small amplitude[4] and at large amplitude[5]. This paper describes our group's implementation of a new torsion balance method: the measurement of the second harmonic response of the oscillator. This method turns out to be much less sensitive to slow variations in fiber temperature, a limiting factor in some recent experiments.

2. Quick Overview of Oscillating Pendulum Methods

To avoid confusion, let's define the various angles which arise in the discussion of large amplitude oscillation methods. Let θ be the azimuthal position of the torsion pendulum in the lab frame, and let ϕ be the equilibrium position of this oscillator (adjusted by rotating the fiber support point). Then $U(\theta)=\kappa(\theta-\phi)^2/2$, and the nominal motion of the oscillator is $\theta(t)=\phi+Ae^{-\gamma t}\sin\left[\omega\left(t-t_{eq}\right)\right]$. The existence of an external potential, $U_{ext}(\theta)=\lambda\kappa\cos(\theta-\psi)$, can be detected through deviations from this nominal motion. Summary of our experimental method:

1. Move the pendulum support to various positions, ϕ, and at each position record several periods of free oscillation motion, $\theta(t)$.

2. At each ϕ, fit θ vs. t from the raw crossing time data and determine the value of the signal parameter (old frequency method, or new second harmonic amplitude method).

3. Note the fitted value of signal parameter vs. ϕ, and from the sinusoidal variation in signal parameter, determine λ and ψ.

Now, let's be more mathematically explicit about the response to an external potential, $U_{ext}(\theta)=\lambda\kappa\cos(\theta-\psi)$. For simplicity, ignore damping (set $\gamma=0$) and define $t=0$ as one of the oscillator turning points, then the periodic motion can be expanded in Fourier components as

$$\theta(t)=\phi+A\cos\left[(\omega_o+\Delta\omega)t\right]+C+\sum_{n=2}^{\infty}A_n\cos\left[n(\omega_o+\Delta\omega)t\right]. \qquad (1)$$

For weak external potentials ($\lambda << 1$), the above coefficients can be solved to first-order in λ

$$C=\lambda\sin(\phi-\psi)J_0(A) \qquad \Delta\omega/\omega_o=-\lambda\cos(\phi-\psi)J_1(A)/A$$

$$A_2 = 2\lambda \sin(\phi - \psi) J_2(A)/3 \qquad A_3 = -2\lambda \cos(\phi - \psi) J_3(A)/8$$
$$A_{n\,even} = 2(-1)^{(n+2)/2} \lambda \sin(\phi - \psi) J_n(A) \Big/ \left(n^2 - 1\right)$$
$$A_{n\,odd} = 2(-1)^{(n+1)/2} \lambda \cos(\phi - \psi) J_n(A) \Big/ \left(n^2 - 1\right), \qquad (2)$$

where $J_n(A)$ are Bessel functions, A in radians. Because each of the above depend sinusoidally on $\phi - \psi$, any one of them can be used to detect the external torque, and the strength of that signal has a characteristic dependence on the oscillation amplitude, A.

In particular, we are interested in the use of C, $\Delta\omega$, and A_2 as signals. The magnitude of these signal vs. A are plotted in Fig. 1. For the traditional deflection method (C as signal), the maximum signal occurs at 0° amplitude, and there is no reason to use an oscillating pendulum. For the frequency method, signal-to-noise considerations (discussed below) define an optimal amplitude of 105°. For our new second harmonic method, the maximum signal-to-noise occurs at 175°, but in order to get a sufficient number of crossings to fit for all relevant parameters, it turns out we must operate at 230° amplitude.

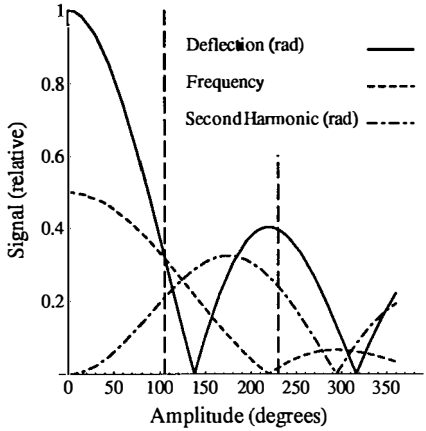


Fig. 1. Signal magnitude (relative) vs. oscillation amplitude (degrees)

When using the frequency method, there is a large penalty in signal-to-noise for operating at small amplitude. The errors in the torsion frequency estimate scale as $1/A$ for most noise processes (e.g., additive measurement noise or thermal noise). Thus, although the $\Delta\omega/\omega_o$ signal is proportional to $J_1(A)/A$, the signal-to-noise varies as $J_1(A)$ which is zero at 0° and a maximum at 105°.

3. Initial Trial Results of Second Harmonic Method

In our present apparatus, we use a visible-light semiconductor laser, an eight-sided edge-reflector mirror directly attached to the torsion pendulum, and a split photodiode to detect crossings. A crossing time (to the nearest 0.5 microsecond) is thus recorded every 45° of pendulum motion. In our second harmonic method, we fit for eight parameters of the equation of motion: amplitude, phase, frequency, damping constant, deflection, linear drift in deflection, second harmonic amplitude (this is the signal parameter), and third harmonic amplitude. In addition, we fit for the angles between the octagonal mirror faces because they are only nominally 45°.

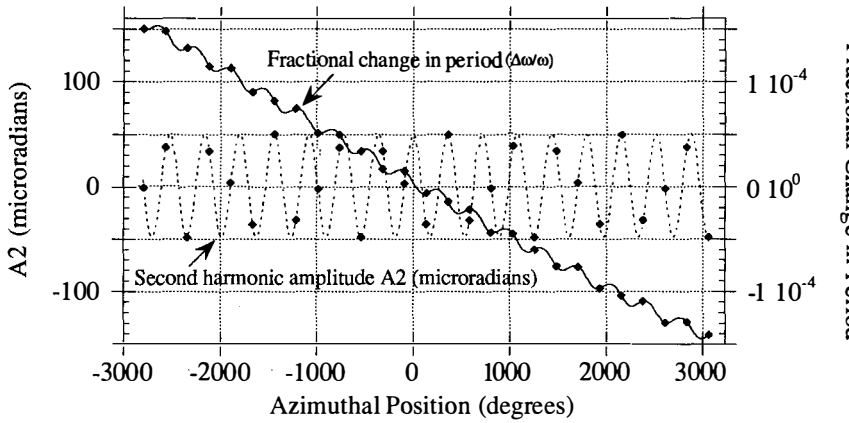


Fig. 2 . Comparison of frequency and second harmonic: exaggerated temperature variation (1.5 °C).

An early test, depicted in Fig. 2, convinced us of the temperature insensitivity of the new method. This 230° amplitude run was taken after a vacuum system bake-out, and so the temperature is decreasing by 1.5 °C over the duration of the run (about 21 hours). The pendulum used had an exaggerated quadrupole mass moment, and the local tidal force (due to a proximate cliff face) was not cancelled, resulting in a very large gravitational gradient signal (signal-to-noise is 5000). It is important to note that the temperature-induced drift in frequency is unmistakable, while the linear drift in the second harmonic signal, A_2, is *at least* 10^{-4} smaller. We don't know exactly how much smaller because the linear drift in A_2 is less than measurement error.

Ultimately, however, we want to apply the method to null experiments with subnanoradian uncertainties, rather than the 50,000 nanoradian signal displayed in Fig. 2. We acquired 3 weeks worth of data using a cylindrically symmetric, single material, null pendulum that by design should produce no signal. Besides the reduction in quadrupole moment, the local gravitational $\cos(\phi - \psi)$ field gradient was compensated with lead plates so that the tidal interaction is a million times smaller than in Fig. 2. Fits from a 16-hour sample of those data are plotted in Fig. 3. The large variation in A_2 is the $\cos(2\phi - \psi_2)$ gravitational gradient from the cliff coupling to machining errors (6 μm tolerance) in the pendulum. This tidal field was not compensated because it is orthogonal to the signal we would be looking for in our null experiments. More troubling is the smaller $\cos(\phi - \psi)$ variation in A_2. This is not gravitational coupling to the outside world because it rotates with the vacuum can (see Fig. 4). We have not definitively identified the cause of this coupling, although it may be a magnetic coupling to the ion pump magnet which does rotate with the vacuum can. The

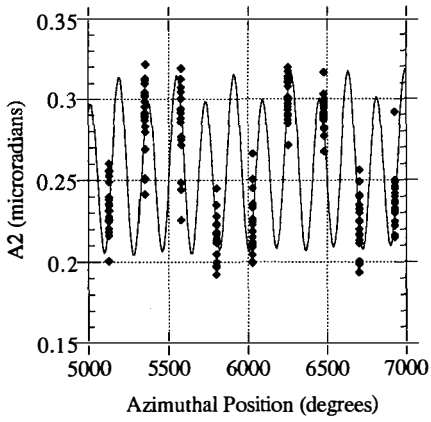


Fig. 3 . Sample null pendulum trial data.

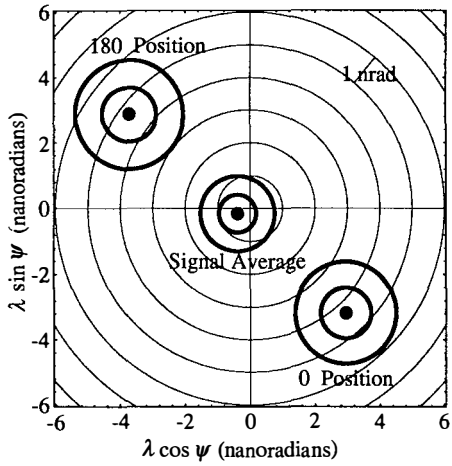


Fig. 4. Null pendulum trial results (3 weeks).

constant term in A_2 is caused by the non-linearity of the torsion fiber. The noise is much less than in Fig. 2, and a small secular linear drift in A_2 can be detected. This is not a temperature dependent effect, but is due to a fiber relaxation process (probably work hardening), which also produces observable drifts in ω, γ, and C.

In Fig. 4, the signals for the 0° vacuum can position, 180° vacuum can position, and their sum for a 3 week run are depicted in the $(\lambda \cos\psi) \otimes (\lambda \sin\psi)$ plane with 1 and 2 sigma confidence contours. Note that the combined result is within 1 sigma of the expected zero signal, and subnanoradian precision has been achieved. We need to investigate the cause of the difference between the 0° and 180° results, but to put these into perspective, the forces which are responsible for this difference are on the order of a nanodyne.

4. Fiber Properties and Applications of Second Harmonic Method

Because of the precision with which we are measuring the torsion pendulum, the effects of non-linearities in the torsion fiber cannot be ignored. The cubic and quartic anharmonic terms produce respectively a second harmonic amplitude and third harmonic amplitude.

A variation in the cubic anharmonic term can possibly produce a false second harmonic signal. At 230° amplitude, we observe a fiber-induced second harmonic of 3×10^{-7} radians. This is too small to detect any temperature dependence, but if the fractional change per °C is comparable to that of the shear modulus, then the temperature induced change in signal is possibly a million times smaller for the second harmonic method than for the frequency method.

Our group is currently applying this new second harmonic method in our experiment at Index, Washington. This experiment uses a magnesium/beryllium pendulum and a granite cliff source mass to test for the presence of composition-dependent forces. This new method will also be employed in our test of the inverse square law of gravity using laboratory source masses. The apparatus for this experiment is currently being built, and we will conduct the experiment in our lab at the University of Washington (Seattle). In the future, we hope to apply this method to even more precise tests of the Weak Equivalence Principle at a site in Hanford, Washington which we are developing with Riley Newman's group at the University of California at Irvine.

References

1. Y. Su et. al., *Phys. Rev. D* **50**, 3614 (1994).
2. P.G. Nelson et. al., *Phys. Rev. D* **42**, 963 (1990).
3. R. Cowsik et. al., *Phys. Rev. Lett.* **64**, 336 (1990).
4. G.G. Luther and W.R. Towler, *Phys. Rev. Lett.* **48**, 121 (1982).
5. P.E. Boynton, in *New and Exotic Phenomena '90, proceedings of the X^{th} Moriond Workshop*, Les Arcs, France, January 1990, eds., O. Fackler and J. Tran Thanh Van (Editions Frontières), p. 207.

Acknowledgements

The authors thank Brent Jones for his help with data analysis and preparation of graphs. This work was supported largely by the National Science Foundation under grant PHY9309468AM02.

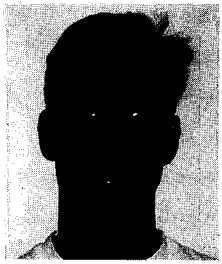
MESURES PRÉCISES DE FORCES FAIBLES

Résumé. Après plus de 200 ans, la balance ou pendule de torsion est encore l'un des moyens les plus précis de mesurer de faibles forces a portée macroscopique. Historiquement, deux principales techniques de mesures ont été developpées pour cet instrument. Dans l'une, le signal est le déplacement a l'equilibre du pendule resultant de la presence d'un couple externe. Dans l'autre, ce couple est détecté par le changement de la fréquence naturelle des oscillations de torsion. Les deux techniques ont été rafinées au point que quelques expériences sont maintenant limitées par les variations temporelles de la température de la fibre de torsion, en depit d'efforts particuliers pour assurer une température stable.

Nous avons developpé et testé une autre technique de mesure a l'aide d'un pendule de torsion, pour laquelle les effets dus aux variations de la température de la fibre sont supprimés par plus de quatre ordres de grandeur. Nous observons l'amplitude de la deuxième harmonique du mouvement du pendule. Une mesure recente sur 21 jours en utilisant un pendule nul (pas de signal attendu) avec Q = 4000 a donné une amplitude de 0.4±0.6 nanoradians pour le deuxième harmonique. Ceci est equivalent pour un pendule d'équilibre a un déplacement de 1.6±2.4 nanoradians. Cette nouvelle technique suggére qu'une amélioration significative dans la précision de ces expériences de physique fondamentale peut être atteinte avec un Q relativement élevé et un appareil a une température d'environ 300K.

Plans for a Redetermination of G

Michael Bantel, R.D Newman, and Z.R. Wang
Department of Physics and Astronomy
University of California Irvine
Irvine CA, 92717 USA
email: mbantel@uci.edu



ABSTRACT

We discuss an experiment under development at UCI to determine the gravitational constant G. A torsion pendulum operating at T≈4K will measure the oscillation frequency shift induced by a pair of ring source masses outside the cryogenic dewar. The source mass rings will produce an extremely pure quadrupole field gradient at the position of the pendulum. The pendulum is to be a small thin quartz plate suspended in a vertical plane by an Al5056 fiber in high vacuum. The measurement will be highly insensitive to uncertainty in the mass distribution and location of the pendulum. Measurements will be made with the pendulum oscillating at each of several large amplitudes (2.57, 4.21, 5.81, ... radians) at which the frequency shift is an extremum as a function of amplitude. This improves the signal/noise ratio relative to small amplitude experiments, makes results insensitive to error in amplitude determination, and offers consistency checks for a variety of systematic effects. The experiment will be performed in a remote former Nike missile bunker which offers very low seismic noise and low time-varying gravitational gradients.

1. Introduction

Four recent precision experiments have measured G, yielding results with assigned uncertainty on the order of 100 ppm which are all mutually exclusive.[1 - 4] These discrepancies have motivated our lab to undertake a measurement of G. In our planned experiment, G is to be determined by measuring the change in oscillation frequency of a thin plate torsion pendulum when a pair of ring shaped source masses (Figure 1) is moved so that their symmetry axis is alternately parallel to the indicated x and y axes. The "dynamic method of measuring G through the effect of source masses on a torsion pendulum's frequency, used most recently by Luther and Towler[4], has the great advantage of avoiding the need for a calibrated reference force. Our pendulum will operate in a cryogenic environment offering several advantages: Thermal noise is greatly reduced (directly through low temperature and indirectly through higher mechanical Q); The pendulum's oscillation frequency stability is greatly increased both by the excellent temperature control possible at low temperature and because the temperature dependence of the pendulum dimensions and fiber shear modulus become small at low temperature; and other fiber properties such as drift and strength improve at low temperature. An advantage of our design is that the pendulum will operate in a relatively large chamber, minimizing pendulum-wall interactions, and will be far from the source mass, reducing sensitivity to source mass inhomogeneity and placement error. The price paid for these advantages is a much smaller signal strength: the fractional frequency shift in our experiment will be more than two orders of magnitude smaller than in the experiment of Luther and Towler[4].

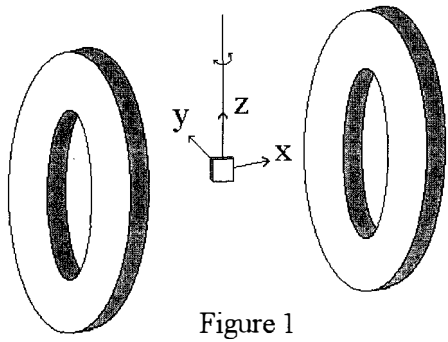


Figure 1

2. Design details

As presently planned, the source masses will be 61 kg OFHC copper rings, with approximate O.D 52 cm, I.D 31 cm, thickness 5 cm. The pendulum is to be a 14 gram fused quartz plate 4 x 4 x 0.4 cm with a reflective coating. The fiber will be aluminum 5056, 30 cm in length, 25 microns in diameter, supporting 50% of its breaking strength. The pendulum period will be about 160 seconds, and the difference in period due to the source ring positions is 6 milliseconds. The operating temperature of the fiber and pendulum will be at either 2K or 4.2 K. The dewar is 46 cm

in diameter, 290 cm in length, and accommodates a 30 cm diameter insert. The rings will hang outside the dewar at room temperature.

3. Analysis formalism and key design features

The gravitational torque N(θ) on a torsion pendulum is given by:

$$N(\theta) = \sum_{lm} i\, m\, q_{lm} a^{*}_{lm} e^{im\theta} \tag{1}$$

where θ is the angular displacement of the pendulum relative to the source masses, the q_{lm} are the mass multipole moments of the pendulum, and the a_{lm} are the field multipole moments of the source masses. The equations for q_{lm} and a_{lm} are:

$$q_{lm} = \int_{pendulum} r^{l}\, \rho_{p}(r)\, Y^{*}_{lm}(\theta,\varphi)\, d^{3}r \qquad a_{lm} = -\frac{4\pi G}{2l+1} \int_{source} \frac{\rho_{s}(r)}{r^{l+1}} Y^{*}_{lm}(\theta,\varphi)\, d^{3}r \tag{2}$$

for a coordinate system centered on the pendulum. The gravitational torque shifts the natural oscillation frequency ω_0^2 of the pendulum:

$$\omega^{2} = \omega_{0}^{2} - \frac{2}{\theta_{0} I} \sum_{l=1}^{\infty} \sum_{m=-l}^{l} m\, q_{lm} a^{*}_{lm} J_{1}(m\theta_{0}) \tag{3}$$

where I is the moment of inertia of the pendulum and θ_0 is the oscillation amplitude. A key feature of our design is that the source mass rings generate a field characterized by a_{lm} which vanish for l=1, 3, 4 and 5; for odd l by symmetry, and for l= 4 as a result of a particular ring spacing. The result is a highly pure quadrupole source field at the pendulum position. The pendulum frequency shift associated with the change in source mass positions is then given to high accuracy by:

$$\Delta\omega^{2} \equiv \omega_{+}^{2} - \omega_{-}^{2} \cong 16\, |a_{22}|\, \frac{|q_{22}|}{I}\, \frac{J_{1}(2\theta_{0})}{\theta_{0}} = KG\, \frac{J_{1}(2\theta_{0})}{\theta_{0}} \tag{4}$$

where K is a parameter based only on the mass distributions of the pendulum and source masses. Corrections to equation (4) from higher order terms are at the level of about 1 ppm, assuming perfect source masses. Important features of our experiment design are:

A. Insensitivity to pendulum dimensions and mass inhomogeneity. In the limit that the pendulum is perfectly planar, the factor |q22|/I in equation (4) is completely independent of the pendulum mass distribution in the plane.

B. Insensitivity to pendulum position. The purity of the quadrupole field is such that a 3 mm error in pendulum position would produce an error in G of less than 1 ppm.

C. Insensitivity to error in oscillation amplitude determination. Previous measurements of G using the dynamic method have operated at small oscillation amplitude to minimize errors due to uncertainty in θ_0. We will operate at several large amplitudes (2.57, 4.21, 5.81, 7.40, 8.98 radians) at which $\Delta\omega^2$ is an extremum and hence extremely insensitive to uncertainty in θ_0. Optical lever timing signals from the four mirrored side faces of the pendulum allow for an extremely accurate determination of θ_0. Large amplitude operation is also favored by the fact that noise contributions to frequency measurements decrease with amplitude as $1/\theta_0$, thus providing a larger signal to noise ratio.

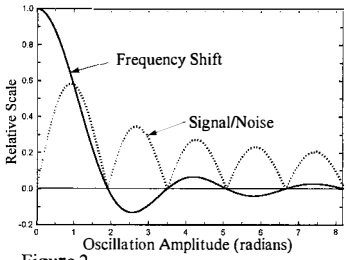


Figure 2

D. Consistency checks. Comparison of G determinations made at widely different oscillation amplitudes affords a powerful check for a number of possible systematic effects.

4. Metrology error sensitivity studies

Evaluating the integrals of equations (2) numerically through $l = 8$ and using equation (3), errors in G due to many assumed deviations from ideal geometry and mass distribution have been calculated. Each such calculation requires about 5 minutes, compared to about 11 hours for a six dimensional point to point torque calculation of comparable accuracy. The dominant error contribution is uncertainty in mean ring separation; a pair of fused quartz rods will keep the rings separated at a known distance with about one micron uncertainty, corresponding to a 2 ppm G error.

5. Fiber nonlinearity, anelasticity and background fields

An equation of motion for the pendulum, which includes possible fiber nonlinearities and background torques resulting from ambient fixed gravitational gradients or electromagnetic couplings, may be written as:

$$I\ddot{\theta} = -k_1\theta - k_2\theta^2 - k_3\theta^3 \pm 4|a_{22}||q_{22}|\,Sin(2\theta) - \sum_m \left[\alpha_m Cos(m\theta) + \beta_m Sin(m\theta)\right] \quad (5)$$

where $\pm 4|a_{22}||q_{22}|$ Sin(2θ) is the signal torque from the source rings. The oscillation frequency of the pendulum may be given as a series expansion to second order in the small torque parameters:

$$\omega^2 = \omega_0^2\left[1 + \sum_i C_i(\theta_0)u_i + \sum_{i,j} D_{i,j}(\theta_0)u_i u_j\right] \quad (6)$$

where the u_i are ratios of the small parameters to k_1 (e.g. k_2/k_1, α_6/k_1,...). We have calculated the first order coefficients $C_i(\theta_0)$ in equation (6) analytically and the second order coefficients $D_{ij}(\theta_0)$ numerically to high precision. Rotating the source mass by 90° around the fiber axis flips the sign of only the signal term in equation (5), so the resulting frequency change is:

$$\Delta\omega^2 = \omega_0^2\left[2C_s(\theta_0)u_s + 2\sum_{s,j} D_{s,j}(\theta_0)u_s u_j\right] \tag{7}$$

where the subscript s represents the source parameter $4|a_{22}||q_{22}|$ Sin(2θ). The error introduced in the G measurement by the background parameters is thus only at second order, given by:

$$\frac{\delta G}{G} = \frac{\delta(\Delta\omega^2)}{\Delta\omega^2} = \frac{2\sum_j D_{s,j}(\theta_0)u_j}{C_s(\theta_0)} \tag{8}$$

The contributions to this error from chamber-fixed and lab-fixed fields are expected to be very small, and may in principle be eliminated by repeating measurements after rotating the pendulum support 90° relative to its chamber and/or rotating the whole apparatus 90° relative to the lab.

Kuroda[5] has pointed out a potentially serious bias in G measurements of this kind, associated with anelastic loss mechanisms in fiber materials. Elsewhere in these proceedings[6] we consider the implications for our measurements of such anelastic effects and also of fiber nonlinearities characterized by k_2 and k_3 in equation (5). Based on our measurements of fiber properties at low temperature, we conclude that these effects should not affect our G measurement by more than 2 ppm.

6. Experiment site

After development at UCI, the apparatus will be operated at an abandoned underground Nike missile site in Hanford, Washington, which is shared between UCI and Professor Paul Boynton at the University of Washington. This isolated site is several kilometers from any cultural activity, and experiences very low seismic noise and time varying gravitational gradients. Our experience with cryogenic torsion balances at UCI indicates that the dominant noise source is seismic (we observe a factor of 20 reduction in frequency measurements error when the entire experiment is suspended using garage door springs). Measured seismic noise power above about 1 Hz at the Hanford site is two orders of magnitude lower than in our area.

7. Anticipated accuracy

We expect to measure G to about 10 ppm, with accuracy limited roughly equally by metrology and temperature variations. If seismic and readout noise prove to be as low as we anticipate, it should be possible to push the measurement to between 1 to 5 ppm accuracy using fused silica source mass rings which would allow optical metrology techniques, and using improved temperature control techniques. In 1942, P.R. Heyl (who measured G at the U.S National Bureau of Standards in 1930 and again in 1942) wrote "The conclusion may be drawn that the limit of the possibilities of torsion balances has been reached"[6]. We hope a half century of technological advances will allow us to push the torsion balance beyond Heyl's expectations.

References

[1] W. Michaelis, (Physikalisch-Technische Bundesanstalt), April 1995 meeting of the APS, Washington DC.

[2] Hinrich Meyer, (BUGH Wuppertal), April 1995 meeting of the APS, Washington DC.

[3] Mark Fitzgerald, (Measurements Standards Laboratory of New Zealand), April 1995 meeting of the APS, Washington DC.

[4] Gabriel G. Luther and William R. Towler, Phys. Rev. Lett. **48**, 121 (1982)

[5] Kazuaki Kuroda, Phys. Rev. Lett. **75**, 2796 (1995)

[6] R.D. Newman, M. Bantel, and Z.R. Wang, *Prospects for gravitational physics experiments with a cryogenic torsion balance*, Rencontres de Moriond (1996)

[7] Paul R. Heyl and Peter Chrzanowski, J. Res. Nat. Bureau. Stand., 29, 1, (1942)

PLANS POUR UNE REDÉTERMINATION DE G

Résumé

Nous discutons une expérience sous développement à UCI à déterminer la constante de gravitation G. Un pendule de torsion fonctionnant à T ≈ 4K mesura le changement de fréquency d'oscillation induit par une paire de masses de source d'anneau a l'exterieure du dewar cryogenique. Les anneaux de masse de source produiront une très pur pente de champ quadrupole à la position du pendule. Le pendule doit être une petite mince plaque de quartz suspendu dans un plan vertical par un fibre A15056 en vide poussé. Le mesurage sera très insensible à l'incertitude de la distribution de masse et de l'emplacement du pendule. On feront les mesurages avec le pendule oscillant à chaque de plusieurs grandes amplitudes (2.57, 4.21, 5.81, ... radians) auxquelles le changement de fréquence est un extrémum comme fonction d'amplitude. Ceci améliore le rapport de signal/bruit par rapport àux expériencs de petite amplitude, fait les resultats insensibles à erreur en détermination d'amplitude, et presente verifications de consistance pour une variété des effets systématique. L'expérience sera effectuer dans un éloigné ancien Nike missile bunker qui offre trés bas bruit séismique et basses variation en temps des pentes de gravitation.

A ROTATING TORSION BALANCE EXPERIMENT TO MEASURE NEWTONS CONSTANT G

J. H. Gundlach
Physics Department
Nuclear Physics Laboratory, Box 354290
University of Washington, Seattle, Washington 98195

ABSTRACT

A new method for measuring Newton's constant G using a continously rotating torsion balance that is operated in a velocity feedback mode is presented. The proposed method has several conceptually new and important features that reduce sensitivity to the dominant systematic uncertainties of previous experiments. We have successfuly conducted exploratory tests to establish the feasibility of this new technique.

1. Motivation

The Newtonian gravitational constant, G, is the least precisely determined fundamental constant. The accepted CODATA[1] value, $G = (6.67259 \pm 0.00085) \times 10^{-11} m^3 kg^{-1} s^{-2}$, is heavily dominated by the 1982 measurement of Luther and Towler[2] and was assigned an uncertainty of 128 ppm. Recently this value has been brought into question by several groups. The German Physikalisch Technische Bundesanstalt[3] obtained a value 0.6% (~40 standard deviations) higher, a New Zealand group[4] reported a value 0.1% (~7 standard deviations) lower, while a Wuppertal group[5] obtained a value 0.06% lower than the CODATA value. In addition a Russian group[6] claimed to observe a temporal and length-scale variation of G at the 0.7% level. Except for ref. [5] which used a new double-pendulum technique, these experiments employed the classical strategy of measuring the torque on a torsion pendulum or relied on the constancy of the restoring torque of a torsion fiber undergoing large-amplitude oscillations. Kuroda[7] recently pointed out that G measurements based on detecting the change in torsional oscillation frequency may have a systematic bias due to torsion fiber inelasticity. A decisive measurement, preferably using a new technique, is needed to resolve the discrepancies in the value of this natural constant.

2. The Proposed Method

We have developed a new method for measuring G that is based on measuring the *angular acceleration* of a "two dimensional" torsion pendulum. Our method overcomes important sources of systematic error in previous measurements. In particular, the pendulum dimensions, mass, and density distribution need not be known precisely, and many torsion fiber properties need not be known precisely or even remain constant. We have run numerical simulations and have conducted exploratory measurements using an existing apparatus to demonstrate the feasibility of this new method.

The gravitational angular acceleration, α, of a torsion pendulum in the field of a nearby attractor can be expressed in a multipole formalism[8, 9].

$$\alpha(\phi) = \sum_{l,m} \alpha_{l,m} = -\frac{4\pi G}{I} \sum_{l=2}^{\infty} \frac{1}{2l+1} \sum_{m=-l}^{+l} m q_{lm} Q_{lm} e^{im\phi} , \tag{1}$$

where q_{lm} and Q_{lm} are the spherical multipole moments of the pendulum and multipole fields of the attractor respectively, ϕ is the azimuthal angle between the pendulum and the attractor and I is the pendulum moment of inertia. Equation 1 assumes that the restoring torque from the suspension fiber is neglible; we justify this approximation below. With the choice of pendulum and attractor geometries discussed below, α will be dominated by the $q_{22}Q_{22}$ term in Eq. 1. In this case

$$\alpha(\phi) = \alpha_{2,2} = -\frac{16\pi}{5} G \frac{q_{22}}{I} Q_{22} \sin 2\phi . \tag{2}$$

The quotient

$$\frac{q_{22}}{I} = \frac{\int \rho(\vec{r}_p)\ Y_{22}(\theta_p, \phi_p)\ r_p^2\ d^3 r_p}{\int \rho(\vec{r}_p)\ \sin^2\theta_p\ r_p^2\ d^3 r_p} \longrightarrow \sqrt{\frac{15}{32\pi}} , \tag{3}$$

where $\rho(\vec{r}_p)$ is the pendulum density, becomes a constant for a pendulum that lies entirely in a plane that includes the torsion fiber axis, so that

$$\alpha(\phi) = \alpha_{2,2} = -\sqrt{\frac{24\pi}{5}}\ G\ Q_{22} \sin 2\phi . \tag{4}$$

This allows a measurement of α to yield a precise value of G that is independent of the pendulum mass, dimensions or density distribution. It is worth noting that uncertainies in these quantities formed the dominant contributions to the error in the value of ref. [2].

A pendulum and attractor geometry that incorporates these ideas is shown in Fig. 1. A complete torsion pendulum apparatus mounted on a turntable is initially started rotating at a slow rate, ω_t, for example 2π rad/h. Eight massive spheres centered about the pendulum exert a time-varying gravitational torque on the pendulum that, with the geometries discussed below, will be dominated by a $q_{22}Q_{22}$ coupling.

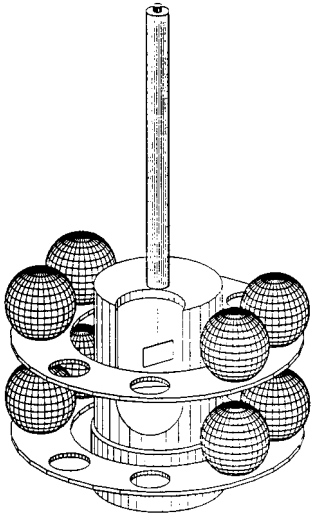


Fig. 1 Diagram of our proposed apparatus for measuring G. The drawing shows only the essential geometrical features. The torsion balance (shown cut away to display the pendulum) sits on a turntable; the rings supporting the spheres are mounted on a second turntable (not shown). The pendulum and attractor spheres are drawn to scale. Asymmetries in the attractor turntable can be cancelled by shifting the spheres to the unoccupied holes in the rings.

The angular acceleration of the pendulum is measured by activating a feedback loop that continually adjusts the rotation rate of the turntable to follow the pendulum so that the torsion fiber never twists from its equilibrium angle. In this way the "free" angular acceleration of the pendulum is directly transferred to the turntable, and the restoring torque from the suspension fiber is driven to zero. Because the fiber never twists, the torsion constant, κ, of the fiber need not be known nor will its inelastic properties directly affect the measurement[12].

The turntable angular acceleration is determined from the change in the pulse rate of a high-resolution angle encoder, a highly linear device that does not require external calibration (because $0° = 360°$). One determines G by fitting $\alpha(\phi)$ with a harmonic series in ϕ and extracting the coefficent of $\sin 2\phi$ to select angular accelerations, $\alpha_{l,m}$, with $m = 2$. To eliminate accelerations caused by other objects in the lab, the attractors are placed on a second turntable that rotates at a rate, ω_a, whose magnitude and sign differ from ω_t. This rotation also averages out any local non-linearities of the pendulum shaft encoder, reduces any effects from vibrations associated with either of the turntables, and puts the signal at a relatively high frequency which reduces noise, in particular low-frequency gravitational noise.

We now discuss techniques for minimizing $l > 2$ gravitational torques so that α is dominated by the $q_{22}Q_{22}$ coupling. The magnitudes of $l > 2$ torques are naturally reduced by factors $(R_p/R_a)^{l-2}$, where R_p is a typical dimension of the pendulum and R_a is the radius to the attractors. The leading higher-l accelerations can be made to vanish with proper pendulum and

attractor design. The q_{l2} and Q_{l2} moments with odd l vanish due to symmetry about the horizontal midplane. A rectangular pendulum with a width w, height h and thickness t will have a vanishing q_{42} moment if $10h^2 = 3(w^2 + t^2)$. By making the attractors from pairs of spheres of mass M, with vertical separation z, at a radial distance from the pendulum axis $\rho = \sqrt{3/2}z$, we eliminate the Q_{42} field; by employing two pairs of spheres on either side of the pendulum, separated by 45° of azimuth, we eliminate all odd m couplings as well as those with $m = 4$. With this design, shown in Fig. 1 we have

$$Q_{22} = \sqrt{\frac{10}{7\pi}}\frac{108}{49}\frac{M}{\rho^3} , \tag{5}$$

while the leading non-$q_{22}Q_{22}$ torque, which occurs in $l = 6$ order, is calculable and small,

$$\frac{\alpha_{6,2}}{\alpha_{2,2}} = \frac{99}{7683200}\frac{213(w^4 + t^4) + 626w^2t^2}{\rho^4} . \tag{6}$$

Note that all lower-order torques, except for the $q_{22}Q_{22}$ torque of interest, are the products of two small (nominally zero) values.

A real world rectangular pendulum with finite thickness has

$$\frac{q_{22}}{I} = \frac{w^2 - t^2}{w^2 + t^2}\sqrt{\frac{15}{32\pi}} . \tag{7}$$

For a pendulum with a width of $w = 76.00$ mm, a height of $h = 41.65$ mm and a thickness of $t = 2.50$ mm, q_{22}/I deviates from the "two-dimensional" value by only 0.2% (q_{42} vanishes) and the dependence of G on the dimensional properties is still weak: One can do a 10^{-5} measurement if w is uncertain by 0.20 mm, or if density variations are as large as 0.46%[10], or if the absolute thickness and overall flatness of the pendulum are uncertain by 5 μm. Similarly, the tip of the pendulum about a horizontal axis can be as large as 2 mrad, or a rotational misalignment about an axis perpendicular to the plane can be as large as $\leq$10 mrad. If the pendulum is fabricated from a quartz glass plate, one can use optical methods to measure its thickness, flatness, and density uniformity. The quartz plate faces can then be Au-coated to reflect the light beam that monitors the angular deflection (and to electrically ground the pendulum). The light beam can also be used to measure the pendulum tip by rotating the pendulum so that the light hits the opposite face of the pendulum. The attractor mass will be constructed from spheres to eliminate problems with density nonuniformities by changing the orientations of the individual spheres. Spherically symmetric density variations do not affect the fields.

A feedback circuit with finite gain requires a small twist of the torsion fiber to derive its feedback signal. Hence a small torque is tied up in twisting the torsion fiber. As the twist angle is recorded, it is straightforward to account for this small extra torque when extracting G from $\alpha(\phi)$. If the open-loop gain (the factor by which the feedback reduces the pendulum deflection) is sufficiently high ($> 10^3$), and the damping time of the free pendulum sufficiently long ($\tau_d \geq 10^4$ s), a 10^{-5} measurement of G can be made even if the free oscillation frequency of the torsion oscillator ω_0, and the angular-deflection calibration are known only to 1%. The angular-deflection calibration, ω_0, and τ_d are easily found by turning off the feedback and observing the pendulum response to a programmed step-change in the turntable angular velocity.[8]

3. Initial Tests

The performance of the feedback loop is of central importance in our technique. We have tested a feedback algorithm, first with numerical simulations and then with a torsion balance normally

used for equivalence principle tests[8, 9]. The feedback loop is digital and executed in software. It senses the torsion fiber twist angle θ as measured by an autocollimator and adjusts the frequency, ν, of the oscillator controlling the turntable rotation rate so that the autocollimator signal remains unchanged. This frequency, directly proportional to the turntable rotational velocity $\omega_t = \dot{\phi}$, is updated at regular time intervals, Δ, and recorded along with the values of ϕ and θ. The feedback loop uses differential, direct, and integral terms to compute the frequency change

$$\nu_{i+1} - \nu_i \propto c_1 \frac{\theta_i - \theta_{i-1}}{\Delta} + c_2 \frac{\theta_i + \theta_{i-1}}{2} + c_3 \Delta \sum_{m=0}^{i} \theta_m \ . \tag{8}$$

Stable performance is obtained with $c_3 = \Delta/\bar{\tau}^3$, $c_2 = 3\Delta/\bar{\tau}^2 - \omega_0^2\Delta - 1.5c_3\bar{\tau}$, and $c_1 = 3\Delta/\bar{\tau} - c_2\Delta - c_3\Delta^2$, where $\bar{\tau}$ is a characteristic time which should be several times Δ[11].

Figure 2 shows a numerical simulation of a G measurement using the pendulum dimensions given above, an attractor with $Q_{22} = 4.6$ g/cm^3, realistic values for ω_0, and τ_d. The feedback loop had an update time of $\Delta = 2.0$ s. A 3.0 s, low-pass, 6 db/octave analog filter was placed on the autocollimator output, a 1μrad/h fiber drift rate and $1/f^2$ noise in θ, consistent with that observed in previous experiments[9], was included. The average turntable angular velocity was $\bar{\omega}_t = 4.87$ mrad/s, which gave a 4.9% gravitational speed modulation. The open-loop gain could be made to exceed several times 10^3. The simulation gave a G value accurate to 10^{-5} in the equivalent of one day of operation.

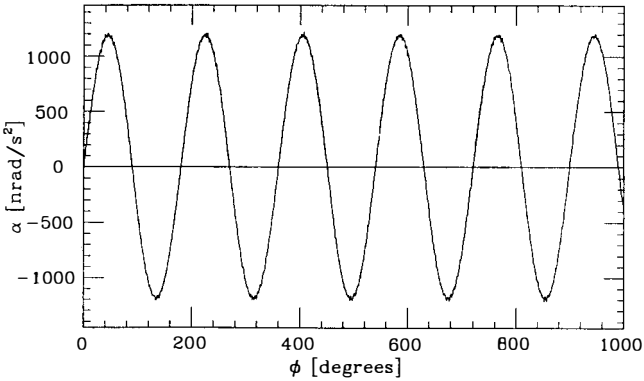


Fig. 2 Numerical simulation of a G measurement, showing the turntable angular acceleration for the pendulum and attractor geometries discussed in the text.

We then successfully implemented the same feedback scheme in an existing equivalence principle apparatus using similar parameters.

Figure 3 shows one result from a series of tests of the feedback algorithm using our Eöt-Wash rotating torsion balance[8, 9]. For this test, we used a lab-fixed Pb attractor with $Q_{22} = 0.52$ g/cm^3 (roughly a factor of 10 smaller than would be used in an actual G measurement), and two of the four test bodies of our normal pendulum were removed to create a sizeable q_{22} moment. The average turntable speed was $\bar{\omega}_t = 0.0011$ rad/s, with a 3.7% gravitational $\sin 2\phi$ speed variation, and the feedback gain was $\approx$2000. The extracted value of G agreed with the standard value to within the 2% uncertainty in Q_{22}. We found that gravity gradient fluctuations caused by human activity were the biggest source of noise. These would be considerably reduced by

operating in a more favorable location.

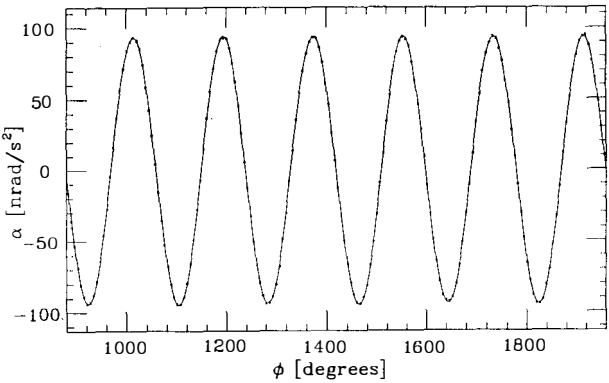


Fig. 3 Proof-of-principle demonstration using the Eöt-Wash torsion balance and a stationary attractor. We show the angular acceleration (averaged over 100 s) of the fedback turntable arising from the predominant $q_{22}Q_{22}$ gravitational coupling. The smooth curve is a harmonic fit to the data. Gravitational fluctuations from human activity in the vicinity of the test setup were the dominant noise source.

4. Conclusion

We have presented a new method for determining G that we believe could provide a substantially improved value good to 10ppm. Our method overcomes the most significant sources of systematic uncertainty encountered in other techniques. Initial tests demonstrate the practicality of the method.

I would like to mention that the input and work from my colleagues Eric Adelberger, Blayne Heckel and Erik Swanson was crucial in the development of this new method.

References

[1] E.R. Cohen and B.N. Taylor, Rev. Mod. Phys., **59** 1121 (1987).
[2] G.G. Luther and W.R. Towler, Phys. Rev. Lett., **48**, 121 (1982).
[3] W. Michaelis, Bull. Am. Phys. Soc. **40**. 976 (1995); results reported in Physics Today **48**, 9 (1995).
[4] M. Fitzgerald, Bull. Am. Phys. Soc. **40**, 975 (1995); results reported in Physics Today **48**, 9 (1995).
[5] H. Meyer, Bull. Am. Phys. Soc. **40**, 975 (1995); results reported in Physics Today **48**, 9 (1995).
[6] V.P. Izmailov, O.V. Karagioz, V.A. Kuznetsov, V.N. Mel'nikov, and A.E. Roslyakov, Measurement Techniques **36**, 1065 (1993).
[7] Kazuaki Kuroda, Phys. Rev. Lett. **75**, 2796 (1995).
[8] E.G. Adelberger, C.W. Stubbs, B.R. Heckel. Y. Su, H.E. Swanson, G. Smith, J.H. Gundlach, and W.F. Rogers, Phys. Rev. D, **42**, 3267 (1990).
[9] Y. Su, B.R. Heckel, E.G. Adelberger, J.H. Gundlach, M. Harris, G.L. Smith, and H.E. Swanson, Phys. Rev. D **50**, 3614 (1994).
[10] An alternative to the feedback scheme discussed here may offer advantages in signal-to-noise. In this scheme the turntable is driven in a smooth, preprogrammed way that closely anticipates the pendulum acceleration.
[11] This value occurs in a worst-case scenario where the density of the central half of the pendulum is 0.46% greater than the density of the outer half.
[12] A small correction for the finite open-loop gain must account for the phase delay and attenuation of the autocollimator read-out.

SPECIAL LECTURES

TEN YEARS OF THE FIFTH FORCE*

Ephraim Fischbach and Carrick Talmadge
Physics Department, Purdue University, West Lafayette, IN 47907-1396 USA

ABSTRACT

The suggestion in 1986 of a possible gravity-like "fifth" fundamental force renewed interest in the question of whether new macroscopic forces are present in nature. Such forces are predicted in many theories which unify gravity with the other known forces, and their presence can be detected by searching for apparent deviations from the predictions of Newtonian gravity. We review the phenomenology behind searches for a "fifth force", and present a summary of the existing experimental constraints.

*Work supported in part by the U.S. Department of Energy.

This year marks the 10th anniversary of the "fifth force" hypothesis — the suggestion that there exists in nature a new intermediate range force similar to gravity, and co-existing with it [1-9]. Much of the work carried out during this period has been reported at Moriond, and so it is appropriate to use the occasion of this year's Moriond Workshop to review what we have learned during the past decade.

In the simplest models, a "fifth force" would arise from the exchange of a new ultra-light boson which couples to ordinary matter with a strength comparable to gravity. There are numerous theories of physics at the Planck scale which predict the existence of such ultra-light bosonic fields [3-6], whose effect is to modify the expression for the interaction energy $V(r)$ for two point masses m_1 and m_2:

$$V(r) = \frac{-G_\infty m_1 m_2}{r}(1 + \alpha e^{-r/\lambda}). \tag{1}$$

Here $r = |\vec{r}_1 - \vec{r}_2|$ is the separation of the masses, and G_∞ is the Newtonian gravitational constant for $r \to \infty$. The constants α and λ characterize the strength of the new interaction (relative to gravity), and the range of the new force. Differentiating $V(r)$ leads to the following expression for the force $\vec{F}(r)$, which is what is measured in most experiments:

$$\begin{aligned} \vec{F}(r) &= -\vec{\nabla}V(r) = -G(r)\frac{m_1 m_2 \hat{r}}{r^2}, \\ G(r) &= G_\infty[1 + \alpha(1 + r/\lambda)e^{-r/\lambda}]; G_o \equiv G_\infty(1 + \alpha), \end{aligned} \tag{2}$$

We see from Eq.(2) that in the presence of a "fifth force" ($\alpha \neq 0$) the usual inverse-square law breaks down. It follows that a search for deviations from the inverse-square law can be interpreted as a probe for new forces, and hence of physics at the Planck scale. The results of any test of the inverse-square law can then be expressed in terms of an exclusion plot in the $\alpha - \lambda$ plane, as shown in Fig. 1. (In anticipation of the ensuing discussion, we note that tests of the inverse-square law are also referred to as "composition-independent" tests for new forces.)

The stimulation for the fifth force hypothesis in 1986 came in part from the recognition that in many specific theories the parameter α in Eq.(1) is not a fundamental constant of nature, but depends on the chemical compositions of the test masses. To understand how this comes about we consider the coupling of new bosonic field to the baryon number $B = N + Z$, where N and Z denote the numbers of neutrons and protons respectively. The additional potential energy $V_5(r)$ arising from the interaction of masses 1 and 2 is

$$V_5(r) = f^2 \frac{B_1 B_2}{r} e^{-r/\lambda}, \tag{3}$$

where f is a new fundamental constant. It is straightforward to show that the sum of Eq.(3) and the usual Newtonian potential leads to Eq.(1) with α replaced by α_{12},

$$\alpha_{12} = -\xi(B_1/\mu_1)(B_2/\mu_2), \tag{4}$$

where $\xi = f^2/G_\infty m_H^2$, $\mu_{1,2} = m_{1,2}/m_H$, and $m_H = m({}_1H^1)$. It follows from Eqs.(1)-(4) that the acceleration difference $\Delta\vec{a}_{12}$ of 1 and 2 towards the Earth is given by

$$\Delta\vec{a}_{12} = \xi(B/\mu)_\oplus[(B/\mu)_1 - (B/\mu)_2]\vec{\mathcal{F}}, \tag{5}$$

where $\vec{\mathcal{F}}$ is the field strength of the source (in units of acceleration), which in this case is the Earth (denoted by $\oplus$). For a coupling to another charge Q, e.g. isospin $Q = I_z = N - Z$, one merely substitutes $B \to Q$ in Eq.(5). Since $\vec{\mathcal{F}}$ depends explicitly on λ it follows that an experimental limit on $\Delta\vec{a}_{12}$ leads to a constraint among the parameters ξ, λ, and Q. In practice the constraints in the $\xi - \lambda$ plane are usually plotted for different choices of Q, as in Fig. 2 for $Q = B$ and $Q = I_z$.

Figures 1 and 2 respectively give the current (as of March 1996) constraints on composition-independent and composition-dependent deviations from Newtonian gravity. In each figure the shading denotes the regions in the $\alpha - \lambda$ or $\xi - \lambda$ plane which are excluded by the data at the 2σ level. We note that in each graph, the lower boundary of the shaded region is determined by superimposing the results of a number of different experiments. As we discuss in Ref. [9], composition-independent experiments achieve their maximum sensitivity for values of λ comparable to the dimensions of the apparatus, and hence no single experiment can be sensitive to all values of λ. The situation is somewhat different for composition-dependent experiments but, for different reasons, it is again necessary to rely on a collection of experiments over different distance scales [9].

One can summarize the current experimental situation as follows: There is at present no compelling experimental evidence for any deviation from the predictions of Newtonian gravity in either composition-independent or composition-dependent experiments. Although there are some anomalous results which remain to be understood, most notably in the original Eötvös experiment [10], the preponderance of the existing experimental data is incompatible with the presence of any new intermediate-range or long-range forces.

We conclude this discussion by briefly summarizing the status of each of the experiments or analyses in which an anomaly was reported.

(1) ***Eötvös, Pekár, and Fekete*** (1922); Ref. [10] The EPF data were the first indication of a possible intermediate-range composition-dependent "fifth force". More recent experiments with much higher sensitivity have seen no evidence for such a force, and hence (by implication) suggest that the EPF results are wrong. However, attempts to find significant flaws in their experiment have failed, as have efforts to explain the EPF data in terms of conventional physics. There remains a slight possibility that by virtue of of its configuration and/or its location, the EPF experiment might have been sensitive to a new force to which other experiments were not.

In any case, the origin and interpretation of the EPF results remain a mystery at the present time.

(2) *Long* (1976); Ref. [11] This work was the motivation for the very careful laboratory experiments of Newman and collaborators, as well as other groups (see, for example, Ref. [12]). None of the more recent experiments confirm Long's results. Subsequent analysis by Long himself suggests that he may have been seeing the effects of a tilt of the floor in his laboratory as his test masses were moved.

(3) *Stacey and Tuck* (1981); Ref. [13] This revival of the Airy method for measuring G_0/G_∞ by geophysical means initially found a result higher than the conventional laboratory value for G_0. Following the analysis of terrain bias by Bartlett and Tew [14], Stacey *et al.* re-examined their data and concluded that the discrepancy between their value of G_∞ and G_0 was a consequence of having undersampled the local gravity field at higher elevations.

(4) *Aronson, et al.* (1982); Ref. [15] This analysis of earlier Fermilab data on kaon regeneration presented evidence for an anomalous energy-dependence of the kaon parameters, such as could arise from an external hypercharge field. Since the effects reported in Ref. [15] have not been seen in subsequent experiments, we are led to conclude that the original data were probably biased by some unknown (but conventional) systematic effect. There is, however, a possibility that these results are correct, notwithstanding the later experiments. This arises from the circumstance that the data came from experiments (E-82 and E-425) in which the kaon beam was not horizontal, but entered the ground at a laboratory angle $\theta_L = 8.25 \times 10^{-3}$rad (to a detector located below ground level). It is straightforward to show that θ_L is related to the angle θ_K seen by the kaons in their proper frame by

$$\tan\theta_K = \gamma \tan\theta_L, \tag{6}$$

where $\gamma = E_K/m_K$ is the usual relativistic factor. For a typical kaon momentum in those experiments, $p_K = 70GeV/c$, $\gamma \cong 140$ and hence $\theta_K \cong 49°$. It follows that the incident kaons in these experiments would have had a large component of momentum *perpendicular* to the Earth, which would not have been the case for the subsequent kaon experiments. It can be shown that motion of a kaon beam perpendicular to a source of a hypercharge field can induce an additional γ-dependence in the kaon parameters [16]. It is thus theoretically possible that the ABCF results are not in conflict with the subsequent experiments, and this could be checked in a number of obvious ways. Similar observations have been made independently by Chardin.

(5) *Thieberger* (1987); Ref. [17] In this experiment a hollow copper sphere floating in a tank of water was observed to move in a direction roughly perpendicular to the face of a cliff on which the apparatus was situated. Although the reported results were compatible with the

original fifth force hypothesis, the results of more sensitive torsion balance experiments carried out subsequently were not. As in the case of the original Eötvös experiment, the implication is that Thieberger's observations can be explained in terms of conventional physics, e.g., as a convection effect.

(6) *Hsui*, (1987); Ref. [18]. This is another determination G_0/G_∞ using the Airy method, based on earlier data from a borehole in Michigan. Since the original measurements were not taken with the present objectives in mind, it is likely that this determination of G_0/G_∞ suffered from the same terrain bias that Stacey, *et al.* encountered. Moreover, a far more serious problem in Hsui's analysis was the imprecise and very limited knowledge of the mass distribution in the region surrounding the borehole, which the author himself noted.

(7) *Boynton, et al.* (1987); [Ref. 19] This torsion balance experiment detected a dependence of the oscillation frequency of a composition-dipole pendant on the orientation of the dipole relative to a cliff. A subsequent repetition of this experiment by the authors using an improved pendant and apparatus saw no effect. Despite efforts to shield the apparatus from stray magnetic fields, it is likely that the original effect was due to a small magnetic impurity in the pendant which coupled to a residual magnetic field.

(8) *Eckhardt, et al.* (1988); [Ref. 20] This was the original WTVD tower experiment in North Carolina which saw evidence for an attractive ("sixth") force. The analysis of terrain bias by Bartlett and Tew [14] suggested that Eckhardt, *et al.*, may have undersampled the local gravity field in low-lying regions surrounding their tower. When the tower results were corrected for this effect, the predicted and observed gravitational accelerations on the tower agreed to within errors. A subsequent experiment by these authors on the WABG tower in Mississippi [21] found agreement with Newtonian gravity, as did experiments on the Erie tower in Colorado [22] and the BREN tower in Nevada [23].

(9) *Ander, et al.* (1989); [24] This was another version of the Airy method, which used a borehole in the Greenland icecap, and observed an anomalous gravity gradient down the borehole. However, this effect could not be attributed unambiguously to a deviation from Newtonian gravity, since it could have also arisen from unexpected mass concentrations in the rock below the ice.

REFERENCES

1. E. Fischbach, D. Sudarsky, A. Szafer, C. Talmadge and S.H. Aronson, Phys. Rev. Lett. **56**, 3 (1986); Ann. Phys. (NY) **182**, 1 (1988).

2. E.G. Adelberger, B.R. Heckel, C.W. Stubbs, and W.F. Rogers, Annu. Rev. Nucl. and Part. Sci. **41**, 269 (1991).

3. Y. Fujii, Int. J. Mod. Phys. **A6**, 3505 (1991).

4. E. Fischbach and C. Talmadge, Nature **356**, 207 (1992).

5. E. Fischbach, G.T. Gillies, D.E. Krause, J.G. Schwan, and C. Talmadge, Metrologia **29**, 215 (1992).

6. C.M. Will, "Theory and Experiment in Gravitational Physics", Revised Edition, (Cambridge Univ. Press, Cambridge, 1993) p. 341ff.

7. A. Franklin, "The Rise and Fall of the Fifth Force" (American Institute of Physics, New York 1993).

8. I. Ciufolini and J.A. Wheeler, "Gravitation and Inertia" (Princeton Univ. Press, Princeton, 1995) p. 91ff.

9. E. Fischbach and C. Talmadge, "The Search for Non-Newtonian Gravity" (American Institute of Physics, New York) in press.

10. R.v. Eötvös, D. Pekár, and E. Fekete, Ann. Phys. (Leipzig) **68**, 11 (1922).

11. D.R. Long, Nature **260**, 417 (1976).

12. J.K. Hoskins, R.D. Newman, R. Spero, and J. Schultz, Phys. Rev. **D32**, 3084 (1985).

13. F.D. Stacey and G.J. Tuck, Nature **292**, 230 (1981); F.D. Stacey, G.J. Tuck, G.I. Moore, S.C. Holding, B.D. Goodwin, and R. Zhou, Rev. Mod. Phys. **59**, 157 (1987).

14. D.F. Bartlett and W.L. Tew, Phys. Rev. **D40**, 673 (1989), and Phys. Rev. Lett. **63**, 1531 (1989).

15. S.H. Aronson, G.J. Bock, H.Y. Cheng and E. Fischbach, Phys. Rev. Lett. **48**, 1306 (1982).

16. D. Sudarsky, E. Fischbach, C. Talmadge, S.H. Aronson, and H.Y. Cheng, Ann. Phys. (NY) **207**, 103 (1991).

17. P. Thieberger, Phys. Rev. Lett. **58**, 1066 (1987).

18. A.T. Hsui, Science **237**, 881 (1987).

19. P.E. Boynton, D. Crosby, P. Ekstrom, and A. Szumilo, Phys. Rev. Lett. **59**, 1385 (1987).

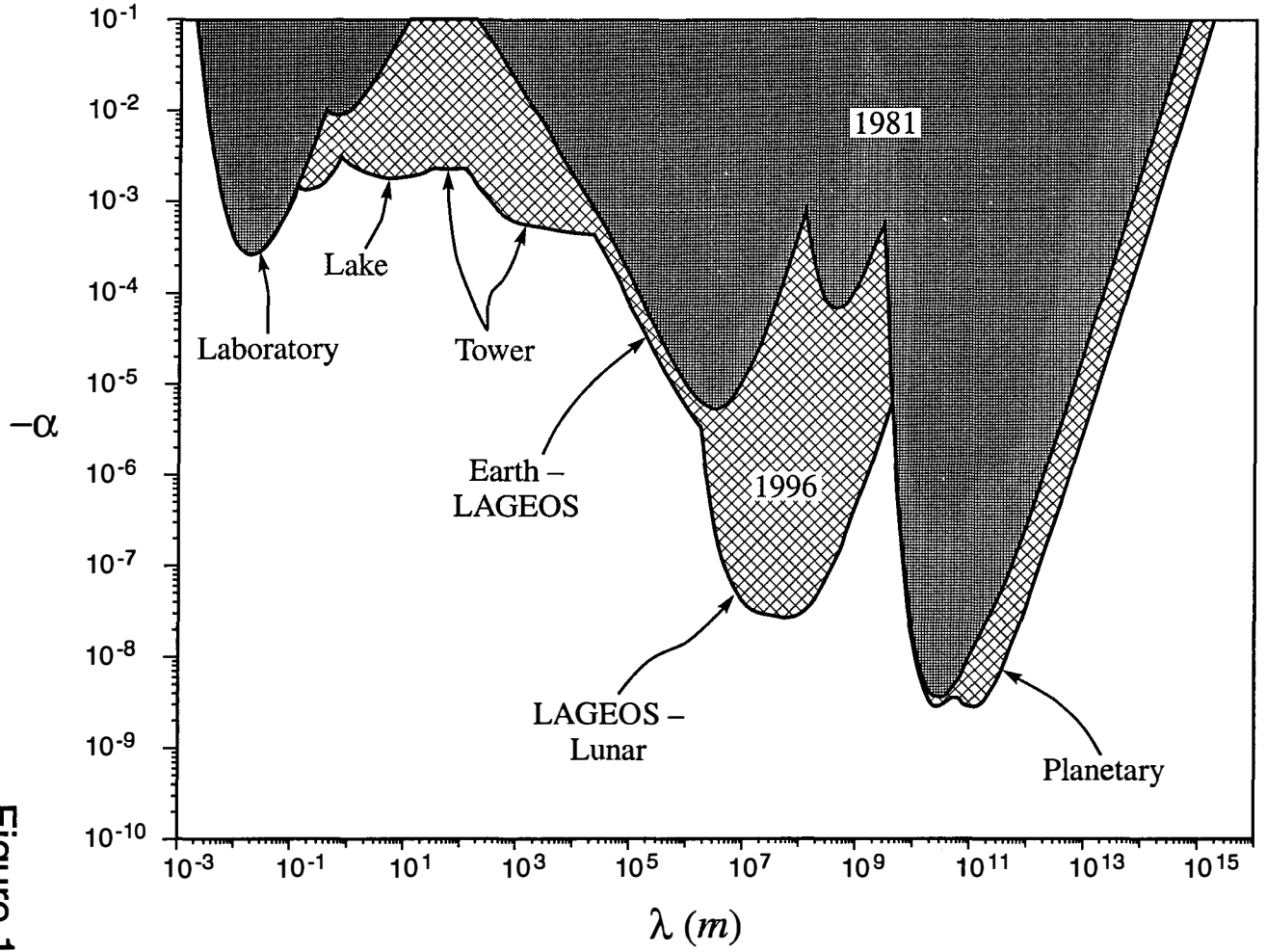
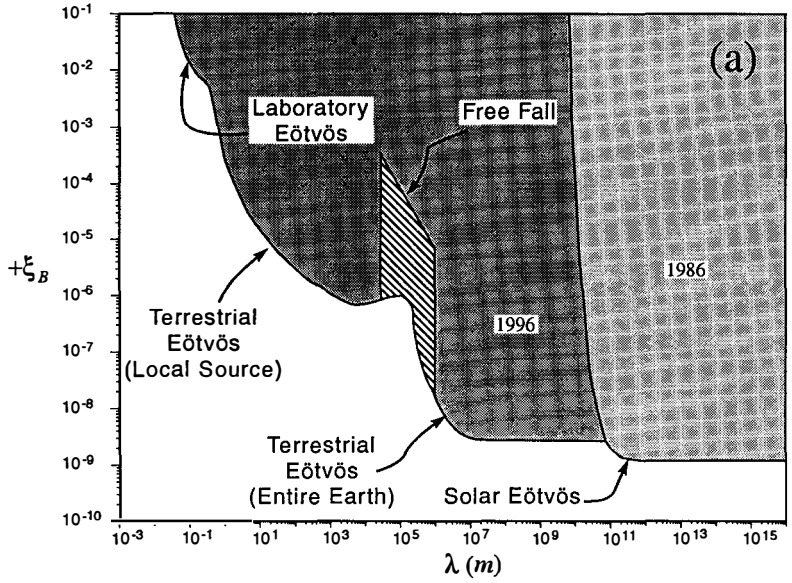

1981
1996
Laboratory
Lake
Tower
Earth –
LAGEOS
LAGEOS –
Lunar
Planetary
−α
λ (m)
10-1
10-2
10-3
10-4
10-5
10-6
10-7
10-8
10-9
10-10
10-3
10-1
101
103
105
107
109
1011
1013
1015

Figure 1


(a)
Laboratory
Eötvös
Free Fall
1986
1996
Terrestrial
Eötvös
(Local Source)
Terrestrial
Eötvös
(Entire Earth)
Solar Eötvös
$+\xi_B$
$\lambda\ (m)$

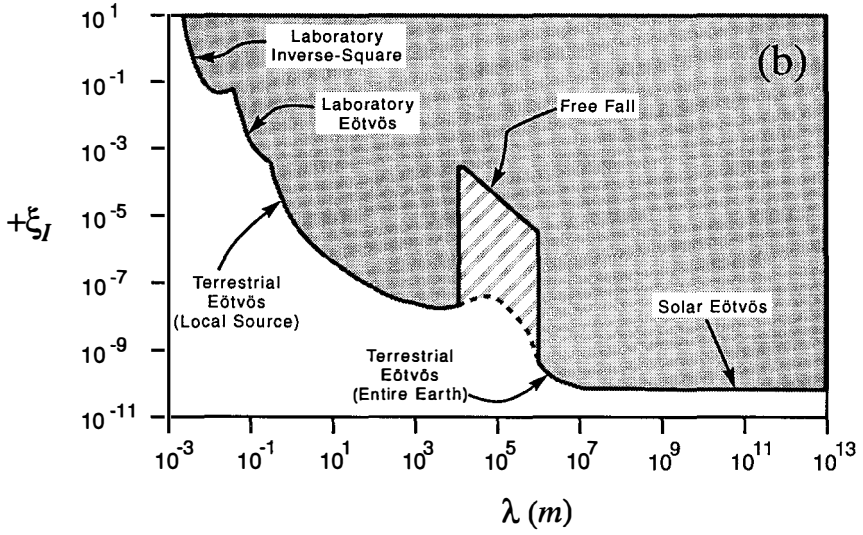

(b)
Laboratory
Inverse-Square
Laboratory
Eötvös
Free Fall
Terrestrial
Eötvös
(Local Source)
Solar Eötvös
Terrestrial
Eötvös
(Entire Earth)
$+\xi_I$
$\lambda\ (m)$

Figure 2

20. D.H. Eckhardt, C. Jekeli, A.R. Lazarewicz, A.J. Romaides, and R.W. Sands, Phys. Rev. Lett. **60**, 2567 (1988).

21. A.J. Romaides, R.W. Sands, D.H. Eckhardt, E. Fischbach, C.L. Talmadge and H.T. Kloor, Phys. Rev. **D50**, 3608 (1994).

22. C.C. Speake, *et al.*, Phys. Rev. Lett. **65**, 1967 (1990).

23. J. Thomas, *et al.*, Phys. Rev. Lett. **63**, 1902 (1989).

24. M. Ander, *et al.*, Phys. Rev. Lett. **62**, 985 (1989).

FIGURE CAPTIONS

Figure 1. Constraints on α and λ in Eq.(1) implied by composition-independent experiments. Results are shown as of 1981 and 1996, and in each case the shaded region is excluded at the 2σ level.

Figure 2. Constraints on $\xi_B(a)$ and $\xi_I(b)$ as a function of λ from composition-dependent experiments. ξ_B and ξ_I are the coupling strengths to $B = N + Z$ and $I_Z = N - Z$ respectively. The shaded regions are excluded at the 2σ level.

Résumé

En 1986 il y avait une suggestion qu'il existait une "cinquieme force" macroscopique dans la nature. Cette idée a stimulé un nouvel intérêt dans cette question. On prédit detelles forces dans beaucoup de theories qui unissent la gravité avec d'autres forces connues. On peut trouver cette présence en cherchant des déviations apparantes des prédictions de la théorie de gravité de Newton. Nous révisons donc la phénomenologie des recherches pour une "cinquieme force", et nous présentons une sommaire des résultats experimentaux courants.

DESPERATELY SEEKING THE SOURCE OF ULTRA HIGH ENERGY COSMIC RAYS

Ludwik M. Celnikier
DAEC, Observatoire de Paris-Meudon, 92 Meudon, France
Celnikier@mesiob.obspm.fr

Cosmic rays

In a century which has witnessed the emergence of several completely new sciences, one would hardly expect a problem originally uncovered over 80 years ago to remain at the cutting edge of contemporary research.

The origin of the highest energy component of cosmic rays is just such a subject (one might note that superconductivity, also discovered in the first decade of the 20th century, is another). The discovery that the Earth is permanently bathed in a flux of ionizing "something" was made in 1912 by the Austrian physicist Victor Hess, who in a series of balloon ascents showed that beyond a certain height the ionization level within a sealed electroscope increases with altitude. The techniques of the time could not identify the nature of this "something"; up to the beginning of the 1930s, it was generally believed that the ionizing flux consisted of high energy gamma rays, and indeed Robert Millikan claimed that the (high) energy spectrum was crossed by "gaps", creating energy "bands" which could be interpreted as the creation energies of massive nuclei (the low energy component of the ionizing flux was clearly related to radioactivity in the Earth's surface). In the absence of any correlation with known celestial sources, high energy cosmic rays could plausibly be ascribed to an early "cosmological" period of element formation; it is just as well to recall that at that time primordial nucleosynthesis was just a dream, and Millikan's speculation was based on little more than Einstein's mass-energy equivalence principle.

In time, the ionizing "something" observed by Millikan's generation was definitely shown to be made of massive particles, essentially the stripped nuclei of the common elements, with a small admixture of electrons; as satellites became available, these particles could be observed directly rather than via the effects of their products on electroscopes, and it became clear that, apart from complications at energies below about 10^{10} eV/particle due to the perturbing effect

of the solar magnetic field, the differential flux j as a function of energy E follows a power law of the type $dj/dE = E^q$, with the index $q \approx -2{\cdot}5$ (but see below).

It was in fact not until the beginning of the 1970s that γ rays were also shown to be a component of the cosmic radiation.

While satellite techniques are admirably suited up to energies of about 10^{14} eV/particle, the low flux beyond $\sim 10^{15}$ eV/particle would require implausibly large detectors or give absurdly low event rates, and a method originally suggested by Pierre Auger comes into its own.

In essence, a high energy particle impinging on the atmosphere interacts and gives rise to secondary particles, which in turn produce other secondaries, and so on. As the resultant "shower" grows in importance, the energy of each secondary decreases until various thresholds are reached. In particular, when the energy of the pion secondaries drops below a few tens of GeV, they can only disintegrate, producing on the order of 1 GeV muons which thereafter reach the ground essentially directly; electrons and gamma rays in the secondaries are however subject to various electro-magnetic scattering processes which act along the entire trajectory and so *their* energy at ground level is typically of the order of 10 MeV. Moreover, as a consequence of their different interaction histories, muons reach the ground before the secondary electrons and gamma rays.

Electrons, γ rays and muons are in principle readily distinguished from each other, for example by suitably instrumented Cerenkov detectors and scintillators, and their staggered arrival constitutes the signature of a cosmic ray event; sampling at ground level with a network of devices allows one to reconstruct the structure and composition of the shower. The total energy of the primary particle is relatively straightforward to deduce from this information, as is the direction of the primary from the inclination of the shower "wavefront"; however, to identify the primary from its final products it is necessary to compute detailed models of the shower as it propagates through the atmosphere. The problem is only a little less difficult than that of finding who owned a smashed watch some of whose fragments have been found lying in the street.

As the shower propagates through the atmosphere it excites a near ultra-violet transition of the nitrogen atom to which a dry atmosphere is relatively transparent, and so the motion of the shower front is associated with a corresponding fluorescent pulse which can be followed using a suitable photoelectric device. It turns out that the intensity of the fluorescent pulse goes through a maximum, at an altitude which depends essentially on the atmospheric structure and the nature of the primary particle. Moreover, the fluorescent intensity at its maximum is a function the energy of the primary. It is amusing to note that there is a qualitative analogy with the Chapman theory of atmospheric ionization by solar ultra-violet and X radiation.

In practise, the variation of the fluorescence with altitude, described by what is usually known as a Gaisser-Hillas function but which has a qualitative resemblance to the Chapman function, is quite complicated, and some subtlety is required to ascertain the nature of the

primary cosmic ray as well as its energy from the fluorescence observations. The technique is nevertheless very powerful and when it can be exploited can identify the primary with much less ambiguity than the air shower arrays. At least two observing stations each equiped with a set of telescopes scanning as much of the sky as possible are required in order to obtain the direction of the primary unambiguously. However, the technique can only work on very clear, cloudless, dark nights, which therefore reduces considerably its efficiency as compared to the conventional giant air shower arrays. The "Fly's Eye" in Utah is at present the only exponent of this technique.

Useful summaries of cosmic ray observing methods are given by Sokolsky[13] and Sokolsky et al.[12].

Ultra high energy cosmic rays

The giant arrays and the Fly's Eye have enabled the cosmic ray spectrum to be extended to well beyond 10^{20} eV/particle – in fact, the most energetic particle so far detected (by the Fly's Eye[4] [7]) had an energy of just over 3×10^{20} eV, which corresponds, be it noted, to just over 50 joules.

A number of intriguing features emerge from all these measurements[5][6].

The energy density of cosmic rays in the region of the Earth is of the order of 1 eV/cm^3; while this is comparable to the energy density of starlight, one should keep in mind that the mean energy per particle is of course very much higher than that per photon.

From about 10^{10} eV/particle (where solar magnetic field effects become negligible) the differential spectrum seems to be a "segmented" power law, with an index $\sim -2{\cdot}7$ up to an energy of about 10^{16} eV, steepening slightly to about -3 up to on the order of 10^{19} eV, after which there is some suggestion that the spectrum becomes somewhat flatter. The spectrum is continuous at around 10^{16} eV, with no discernible discontinuity: unless one invokes some kind of *ad hoc* fine tuning, a single mechanism or two related mechanisms would appear necessary to explain the spectrum over the range $10^{10} - 10^{19}$ eV. However, the flattening of the spectrum at extreme energies is suggestive of the presence of a quite unrelated process.

Up to a few times 10^{18} eV, the cosmic ray chemical composition corresponds to the usual mix of elements, with however an overabundance of the "heavier" component. This latter effect can be understood as a simple consequence of the nuclear reactions which must occur between the "original" cosmic rays and interstellar matter; indeed, the overabundance can be used to infer the typical time cosmic rays spend wandering about before reaching us. However, beyond a few times 10^{18} eV, the chemical composition becomes "lighter"; while at present it is not possible to definitively discriminate against γ rays (but see Halzen et al.[7]) at the highest energies, it would appear that the least unlikely identification of these particles is protons (or technically, but implausibly, neutrinos). And just to add spice to the problem, the very highest

energy event, that observed by the Fly's Eye, could conceivably be due to a heavy nucleus: the shower profile in this case does not provide enough information to discriminate against such a possibility.

There is some suggestion[11] that the spectrum has a gap at a little under 10^{20} eV; if confirmed, this could again imply that a new process sets in at around that energy; however, this statement should be tempered with the knowledge that only about 40 events have been observed beyond 5×10^{19} eV, and 8 events beyond 10^{20} eV.

The sources

Two generic classes of mechanism can produce high energy charged particles in the cosmos without appealing to theology. One is admirably suited to give a power law energy spectrum; the other is an excellent source of very high energy particles, but with no guarantee that their spectrum will follow a power law.

Two criteria must be satisfied for the differential spectrum to have the form of a power law. In essence, the cosmic rays have to be confined to a certain volume of space, from which they extract energy in such a way that the energy increase is a rising function of the particles' energy; a second, quite reasonable, condition, is that beyond some energy the accelerated particles leave the "active" volume.

Many years ago, Enrico Fermi showed that a charged particle "bouncing around" between moving magnetic discontinuities would gain energy from the magnetic field in just the required way. Moreover, once its energy reached a value such that its cyclotron radius is of the order of the characteristic dimension of the confining volume, the particle would necessarily tend to leave.

The simplest way to apply Fermi's theory is to assume energy equipartition between the magnetic field and the randomly moving plasma in which it is embedded, and to farther assume that sufficient time is available for the cosmic ray particles to reach energy equipartition with the moving magnetic "mirrors". Assuming finally that the only way energy is lost from this system is via the evasion of sufficiently energetic particles on a time scale longer than that to reach equipartition, it can be shown very easily[9] that the equilibrium differential spectrum is a power law in energy, with an index exactly equal to $-2{\cdot}5$.

The good news is thus that the observed cosmic ray spectrum could be a simple consequence of a rather general process.

Now for the bad news: close as this result is to observation, it disagrees with it in a fundamental way. As noted in the introduction, the spectrum has at least three different power law indices depending on the energy range. Presumably perfect energy equipartition is not a feature of the acceleration process, and it becomes necessary to compute detailed models. These are parametrised essentially in terms of two quantities:

1. the time scale t_a for the acceleration by the Fermi "bouncing" mechanism, and
2. the time scale t_l to leave the accelerating region, for example by some kind of diffusion process.

In such a case, it can be shown (but not so easily[9]) that the resultant power law spectral index is equal to $-(1 + t_a/t_l)$.

Clearly, different models of the accelerating region lead to different values of the index, and it becomes possible to "fine tune" the result of most models to the observed values... although it must be said that the simplest models of this type lead to power law indices very much farther from observation than the extraordinarily simple, but somewhat miraculous, equipartition theory (this is presumably what miracles are all about!).

Typical configurations which will accelerate particles in this way include[3] the magnetic field of the Sun, the turbulent interstellar medium, shock waves produced by supernova explosions, shock waves produced by relativistically moving compact objects through the interstellar or intergalactic medium, the shocks associated with the motion of galaxies or galaxy clusters through the intergalactic medium, the jets and their endpoints ("lobes" in the jargon of the radio astronomer) associated with certain radio galaxies, and various combinations of such structures.

However, whatever the details of the accelerating process, we know that it can only work if the particle remains trapped in the accelerating region, and it can only remain trapped if its Larmor radius is smaller than the characteristic dimension R of the volume. Therefore, the energy T of a particle of charge Ze is limited by:

$$T < ZeBR \tag{1}$$

where B is the characteristic large scale magnetic field responsible for trapping the particles in the accelerating region.

A totally unrelated mechanism relies on the ability of a suitably constructed homopolar dynamo to produce an intense electric field. Suppose[9] that over some distance scale R a magnetic field B is moving with speed v_B. The induced electric field due to the motion is of the order of $v_B B/c$, and so the energy that a particle of charge Ze can reach is limited by:

$$T < ZeBR\frac{v_B}{c} \tag{2}$$

Clearly, the higher the field and the higher the velocity the better; typical dynamos might be the magnetospheres of neutron stars where magnetic fields could reach on the order of 10^{12} G and the velocities would be relativistic, thereby compensating easily for the small dimensions, and the central engines of active galaxies where it is believed that billion solar mass black holes could be accreting matter and magnetic fields from the surrounding galaxy.

However, while the performance of these "dynamo" accelerators is limited by the magnetic field available (in the first instance, but see later), we know of no natural way for them to generate a power law spectrum. The significance of this depends on one's point of view; we know that the bremmstrahlung radiation of material in the immediate vicinity of pulsars (the "visible" neutron stars) does have a power law spectrum, which can be interpreted as signifying that the pulsars have some way of producing particles with a power law spectrum, even if we do not at present know how.

We do not know *a priori* whether the observed cosmic rays are of galactic origin or not. However, we can impose a distance limit on the source of the highest energy cosmic rays, since the ubiquitous 3 K fossil radiation degrades the energy of particles which pass through it. In effect[9], in the rest frame of, say, a proton with Lorentz factor γ, a photon of the background radiation has energy of the order of $6 \times 10^{-4}\gamma$, and significant pion production can occur if this exceeds 200 MeV. In other words, if the proton energy exceeds about 5×10^{19} eV, it will lose a large part of its energy; the cross-section for this process is known, and leads to the conclusion that no protons with energies significantly in excess of 7×10^{19} can reach us from beyond a few tens of Mpc. This is known as the Greisen-Zatsepin-Kuzmin cutoff: if indeed the ultra high energy cosmic ray component is made of protons, and their source region is beyond this critical distance, the spectrum should stop at $\sim 7 \times 10^{19}$ eV, or should have a gap, the more energetic protons then being produced relatively close by.

The background radiation has yet another effect whose significance in this context is however difficult to assess: ultra energetic heavy nuclei passing through the fossil radiation will be gradually "stripped" down by photodissociation, losing about four nucleons per Mpc. Consequently, a 10^{20} eV proton arriving here could conceivably be the remnant of, say, an iron nucleus which started on its jurney about 50 Mpc away. This phenomenon does increase the volume for potential sources (which nevertheless corresponds, cosmically speaking, to our galactic suburbs); it also suggests that we should not be able to observe heavy nuclei at these energies from such sources. And one would expect this effect to smooth out the Greisen-Zatsepin-Kuzmin cutoff: the detailed topography of the energy spectrum around 10^{20} eV thus carries critical information about the type of process responsible for the ultra high energy particles observed.

Note that analogous effects will operate in compact regions of very high temperature and high photon density – high energy particles will have difficulty leaving such regions "intact".

Finally, γ rays and (of course) neutrinos can reach us from considerable distances – for a proper understanding of the observations, it is vital to eliminate these latter as candidates for the highest energy component.

Equations 1 and 2 are direct functions of the nuclear charge Ze. With a given "technique" it is clearly "easier" to accelerate heavy nuclei to a particular energy than protons – we see in this way the importance of identifying the primary cosmic ray.

These two equations have of course a functional similarity, which allows one to study the

"performance" of the different possibilities in a unified way, by seeing whether the associated magnetic fields, sizes and velocities are capable of creating the highest energy cosmic rays.

It emerges immediately (see Hillas[8] for a very clear summary diagram) that supernova remnant shock waves, turbulent motions within our galaxy, etc. cannot produce protons more energetic than about 10^{16} eV; on the other hand, with a slightly more subtle analysis, it turns out that the energy density of cosmic rays requires only a small fraction of the mechanical energy of supernovae in our galaxy. These latter can thus explain the overall energy budget of cosmic rays, but not the higher energy component. For this, it is necessary to invoke either the shock wave produced by a (hypothetical) outflow of material from our galaxy (such outflows are observed in other galaxies), or the dynamo acceleration of particles by the neutron stars left over from the supernova explosions. The former process would be in fact boosting the lower energy particles to higher energies: consequently, spectral continuity would be maintained right up to the maximum energy that this process can produce, which is nevertheless limited to on the order of 10^{19} eV. The neutron stars could conceivably accelerate particles as far as 10^{20} eV, and since they are the products of the supernovae whose remnants would be responsible for the lower energy cosmic rays, one might intuitively expect some kind of spectral continuity; however, we do not know why the spectrum should be power law, and so for the time being this must be considered an *ad hoc* hypothesis. Furthermore, attractive as neutron stars seem to be as sources of the highest energy cosmic rays, they do suffer from a fundamental defect: the very property which makes them such good candidates – their intense magnetic field – eliminates them, since the accelerated particles will of necessity have to first pass through this very same field, and therefore lose a substantial fraction of their energy through curvature radiation.

The "central engines" of active galaxies are expected to have sizes and magnetic fields capable of accelerating protons to about 10^{20} eV (the expectation emerges essentially from a consideration of the energy available in such objects: the actual processes responsible for the energy conversion, and their efficiency, are far from clear). However, the highest energy protons we are observing (if such they are) cannot come directly from these galactic power houses: we know that the nuclei of active galaxies are sources of intense and high temperature radiation, so that the highest energy cosmic rays will lose most of their energy before leaving the source region as a consequence of photopion production (through essentially the same process as produces the Greisen-Zatsepin-Kuzmin cutoff on the cosmological scale). Of course, there is no reason why the nuclei of active galaxies should accelerate only protons, since the material accreted onto the central black hole must surely contain all the elements: it is not inconceivable that photodissociation by the radiation field around the black hole and by the fossil radiation along the trajectory to us, could leave as a remnant the very high energy protons... but then there should definitely be no ultra energetic heavy elements. And some remarkably fine tuning is required to strip away just the supernumerary nucleons without also degrading the energy of the surviving proton.

We see again the importance of identifying unambiguously the nature of the primary cosmic rays at the highest energies. Moreover, only sources within 30-50 Mpc could contribute significantly through this process.

Rich and compact galactic clusters could accelerate protons to the requisite energies (but this is an affirmation of dubious significance, since at present we only have upper limits on the value of the intergalactic magnetic field), again via shock waves of some kind produced by the individual galaxies. However, we now know that the intergalactic material in such structures is rich in heavy elements – their highest energy products can be exclusively protons only through a process of photodissociation.

Photodissociation of the heavy nuclei will certainly aleviate the problem of their apparent absence at ultra high energies (remembering, however, that the Fly's Eye superevent might have been an iron nucleus)... but at a price. In essence, each nucleon which is knocked off by the cosmological photons carries away its share of the nucleus' energy, so that the total initial energy of, for example, an iron nucleus of which a 10^{20} eV proton survives at Earth, must have been at least $5{\cdot}6 \times 10^{21}$ eV. Now, equations 1 and 2 do suggest that with a given accelerator there might be no fundamental difference between accelerating a proton to a given energy and accelerating a nucleus of charge Z to Z times that energy; however, the number of nucleons in heavy nuclei, and so the required initial energy for a given final proton energy, grows twice as fast as the nuclear charge. Whilst a factor of two may seem like a niggling detail compared to the order of magnitude which separates reality from plausible and calculable theory, one would surely have felt more comfortable had the factor been one half.

Shock waves in the intercluster medium (which is presumably less contaminated by the products of stellar evolution) cannot be eliminated *a priori* as sources for the highest energy cosmic rays even if these are only protons, although one might note that we know even less about the value of magnetic fields on this very large scale than on the intergalactic scale. However, the scale at which this mechanism must operate if it is to produce 10^{20} eV protons with the aid of fields which do not exceed the known upper limits, runs into the hundreds of Mpc range: consequently, even if $> 10^{20}$ eV protons are produced in this way, we will not see them at these energies because of the Greisen-Zatsepin-Kuzmin cutoff.

In fact, the only structures potentially capable of producing $\geq 10^{20}$ eV protons but which have no evident defects of a rather general nature are the jets and shocks ("lobes") associated with certain radio-galaxies.

Desperately seeking

Within 50 Mpc of our galaxy are a handful of galaxy clusters known as the "Local Group". This region also contains a number of powerful radio-galaxies.

A search[14][10][13] for some correlation of the arrival directions of cosmic rays with suitable

structures in the sky has led to the following conclusions.

The distribution is perfectly isotropic up to an energy of about 10^{18} eV. This is to be expected whatever the source region might be: below this energy, the Larmor radius of particles in the magnetic field of the Galaxy is a small fraction of its size, and so the cosmic ray has little chance of preserving any memory of its original direction.

Between about 10^{18} and a few times 10^{19} eV, the distribution is reasonably well correlated with the plane of the Galaxy. This suggests that these particles, and so by extension the lower energy particles, originate from sources associated with our galaxy.

However, the highest energy cosmic rays, up to about 10^{20} eV, seem to be correlated with a plane perpendicular to that of our galaxy. Now, the "Local Group" of galaxies forms a vaguely flattened structure whose plane is in fact perpendicular to the Galaxy, and in which it turns out that many radio-galaxies are to be found. One would be tempted to invoke a "Local Group" origin for these highest energy cosmic rays, were it not for the fact that no specific cases of coincidence have ever been detected. While this is undoubtedly not surprising for cosmic rays whose energy is less than 10^{20} eV, since over tens of Mpc such particles will be deviated significantly from their original directions by the intergalactic magnetic fields (of whose value we only know the upper limit), it is at the very least disturbing for particles well beyond 10^{20} eV: the three highest energy cosmic rays known to date, whose deviations from their original direction cannot exceed ten degrees, come vaguely from the galactic anti-centre direction but have no relation to any known energetic source or galaxy cluster in the local group, and the objects closest to them on the sky are actually at distances which range from 60 to 1000 Mpc, i.e. well beyond the Greisen-Zatsepin-Kuzmin cutoff[10].

Could these ultra high energy cosmic rays be from a completely new type of source, within the Local Group (Greisen–Zatsepin-Kuzmin *oblige*), but in no way related to known galaxies or other structures? Could these particles in fact have been created with such high energies, thereby obviating the need for powerful cosmic accelerators? It is tempting to claim the apparent (but statistically debatable) gap in the differential spectrum below 10^{20} and the possible spectral flattening as indications that something like this is in fact happening[11][10]; massive, but speculative, decaying topological defects which are a possible product of the phase changes which might have followed the Big Bang and which could perhaps have survived to the present epoch have been invoked, as have primordial strings (no less speculative) actually interacting with the Earth's atmosphere... but in the present state of the art there is really nothing very compelling about these *avant garde* theories.

And in the absence of any better idea, there is even some speculation that the source of these particles could be in some way related to γ ray bursters: the feature they share is that we understand neither the one, nor the other.

The shade of Millikan looks on with deep interest and a certain wry smile: his dilemma is still ours, simply moved upwards in energy by a factor of at least 10^{14}, and just as he saw

"bands" in the spectrum, we may (or may not) be seeing a "gap" in the high energy spectrum.

Opening a new astrophysical window

The study of ultra high energy cosmic rays is clearly severely constrained by statistics. In the region of 10^{20} eV the particle flux is of order 10^{-2} events/year/km^2. If the energy spectrum is a continuation of that at lower energies, this event rate drops by two orders of magnitude at 10^{21} eV. With such low fluxes, the essential task of mapping the shape of the high energy spectrum poses a considerable challenge. Moreover, there is at present no operating shower array of any sort in the Southern hemisphere, while even in the Northern hemisphere the sky coverage is very incomplete. This is particularly unfortunate, since the galactic centre can only be observed from the Southern hemisphere, so that the apparent correlations with the anti-centre direction or the supergalactic plane, already statistically weak, could even very easily be a pure artefact. Note too that the recently discovered galactic superluminal sources, which are presumably local, miniaturised, versions of active galactic nuclei and might well be capable of accelerating protons to extreme energies, are all towards the galactic centre.

The "Auger Cosmic Ray Observatory[1]" will be an attempt to improve the observational situation by orders of magnitude.

At the present time, the largest working air shower detector (called AGASA) is in Japan, and consists of about a hundred 2 m^2 scintillators spread over an area of roughly 100 km^2. The only fluorescence detector is the Fly's Eye, in the U.S.A., which consists of two "all sky" photomultiplier stations separated by 3 km; one station has about 900 photomultipliers, and the other a little under 500.

The "Auger Cosmic Ray Observatory" will consist of two "hybrid" detectors, one in the Northern and the other in the Southern hemisphere. Each will:

- have about 1600 water Cerenkov detectors spread over roughly 3000 km^2, with a 1·5 km spacing;
- be associated with two or three (depending on details of the terrain) "fly's eye" type of fluorescence detectors, which will scan the sky with about 50 mirrors and 200 pixels per mirror.

The system is being designed for an operating lifetime of at least twenty years, with minimal maintenance and maximal autonomy. This raises a number of problems, whose solution is being actively sought; two examples suffice to show the nature of the difficulties:

1. the Cerenkov counters must be associated with very fast electronics capable of recognizing short pulses with good discrimination; the electronics must operate on low power and be very reliable;

2. sophisticated triggering techniques are required to avoid contaminating the data with noise; this entails extremely accurate synchronization of the Cerenkovs, and intercommunication between them. The results of a successful trigger will of course have to be communicated to a central station. Since the arrays will be in remote desert locations, and will involve very many detectors, it is quite unrealistic to envisage cable links: radio telemetry will be the watchword.

If all goes well, the system should begin operating in about the year 2000. It is estimated that after ten years in action, the full Auger array should have of order 1000 events above 7×10^{19} eV "in the bag", of which about 500 should be above 10^{20} eV and 5 above 10^{21} eV *if the spectrum continues* with the same gradient as before. Note that up to about 10% of these events will also have been observed by the Fly's Eye detectors.

While there is no guarantee that such an improvement in statistics will necessarily solve the conundrum of the ultra high energy cosmic rays, it is certain that in its absence the field is limited to the wildest speculation.

Contemporary astrophysics is faced by a number of acute problems.

One of them concerns dark matter, which one might (perhaps mischievously) qualify as the study of particles which *should* exist... but until farther notice, don't.

Ultra high energy cosmic rays constitute the inverse problem: particles which *do* exist... but perhaps shouldn't.

References

[1] The Pierre Auger Project Collaboration Design Report. Fermilab, October 1995.

[2] Bhattacharee P. et al., Grand unified theories, topological defects and ultra high energy cosmic rays, Phys. Rev. Lett., **69** (1992), 567.

[3] Biermann P. L., AGN and Galactic sites of cosmic ray origin, p. 217 in in *High energy astrophysics*, ed. J. M. Matthews, World Scientific Press, 1994.

[4] Bird D. J. et al., Detection of a cosmic ray with measured energy well beyond the expected spectral cutoff due to cosmic microwave radiation, Ap. J. **441** (1995), 144.

[5] Drury L. O'C., Acceleration of cosmic rays, Contemp. Phys., **35** (1994),231.

[6] Ginzburg V. L., Astrophysical aspects of cosmic ray research (first 75 years and outlook for the future), Sov. Phys. Usp., **31** (1988), 491.

[7] Halzen F. et al., The highest energy cosmic ray, Astroparticle Physics, **3** (1995), 151.

[8] Hillas A. M., The origin of ultra high energy cosmic rays, p 277 in *High energy astrophysics*, ed. F. K. Lamb, Benjamin/Cummins Publishing Co. Inc., 1985.

[9] Longair M.S., *High energy astrophysics*, Cambridge University Press, 1981.

[10] Sigl G. et al., On the origin of the highest energy cosmic rays, Astroparticle Physics, **2** (1994), 401.

[11] Sigl G. et al., A gap in the highest energy cosmic ray spectrum as a signature of unification scale physics, Science, **270** (1995), 1977.

[12] Sokolsky P. et al., Extremely high energy cosmic rays, Physics Reports, **217** (1992), 224.

[13] Sokolsky P., Cosmic ray observations at extreme energies, p. 157 in *High energy astrophysics*, ed. J. M. Matthews, World Scientific Press, 1994.

[14] Stanev T. et al., Arrival directions of the most energetic cosmic rays, Phys. Rev. Lett., **75** (1995), 3056.

CHERCHE SOURCE DES RAYONS COSMIQUES ULTRA-ENERGETIQUES... DESESPEREMENT

Depuis quelques années, nous savons que le spectre du rayonnement cosmique s'étend en-deça de 10^{20} eV, et nous avons des fortes présomptions qu'il s'agit de protons. Ces particules ne sont associées avec aucune source céleste connue; d'ailleurs, on ne sait pas par quel processus des protons peuvent être accélérés à de telles énergies. Le projet "Auger" vise à mettre l'étude de ce nouveau fenêtre astrophysique sur des bases solides.

Résumés des contributions

I. LA MATIERE SOMBRE

K. Olive - Pourquoi avons-nous besoin de la matière noire non-baryonique ?
Nous passons en revue les observations et les arguments théoriques qui concourent à prouver l'existence de la matière noire.

L. Van Waerbeke - L'observation comme moyen d'investigation de la matière noire
L'amplification gravitationnelle faible des galaxies lointaines permet d'étudier la distribution totale de masse projetée des structures gravitationnelles d'avant plan à toutes les échelles et a été utilisée avec succès pour cartographier la distribution de masse projetée des amas riches à moyen redshift. Cette étude passe en revue les concepts généraux de l'analyse des lentilles. Nous insistons sur la relation existant entre les quantités observables (formes et flux) et les quantités physiques (masse, redshift) et discutons des problèmes d'observation et des développements récents de l'analyse des données qui semblent prometteurs pour une meilleure mesure des signatures de lentilles (distortion et agrandissement) à très grande échelle.

G. Jungman. Etude de la matière noire des particules
Nous présentons une revue des modèles de matière sombre en donnant une importance particulière aux modèles de matière noire froide (axions et neutralinos supersymétriques).

P. Gondolo. Introduction phénoménologique à la détection directe de la matière noire.
La matière noire qui forme le halo de notre galaxie pourrait être constituée de particules interagissant faiblement avec la matière ordinaire : les WIMPs. Malgré le faible taux d'interaction attendu entre ces particules et le cristal d'un détecteur, la détection directe de WIMPs apparaît possible et plusieurs expériences sont en cours actuellement. Nous décrivons, dans cet article, les éléments de base permettant de calculer ces taux d'interaction, en insistant sur les principales sources d'incertitudes théoriques.

L. Mosca. Détection directe de particules candidates pour la «matière sombre»
Les expériences de détection de particules de matière sombre sont brièvement passées en revue. Les principaux problèmes rencontrés dans ce type de recherche sont soulignés et des stratégies pour les surmonter sont envisagées.

V. Chazal. Détection de matière noire avec un bolomètre saphir de 24 g et mesure de l'environnement neutron au laboratoire souterrain de Modane.
La collaboration EDELWEISS se propose de rechercher la matière noire sous forme de WIMPs. En septembre 1994, des premiers résultats on été obtenus dans le laboratoire souterrain de Modane, avec un bolomètre saphir de 24 g. Nous avons mesuré un taux d'événements de 25 evt/kg/keV/j.
Un autre objectif d'EDELWEISS est l'élimination du bruit de fond neutron sur le site expérimental. Nous avons mesuré un flux de neutrons rapides de 2.10^{-7} neut/s/cm^2. Une étude du flux de neutrons thermiques en cours permettra une interprétation globale de l'environnement neutron à Modane.

L. Zerle. Etat de la recherche de CRESST de la matière noire.
Nous préparons l'expérience CRESST qui a pour objet la recherche des WIMPS de la matière noire au moyen de détecteurs cryogéniques équipés de thermomètres au bord de la transition de phase supraconductrice. Dans la première étape de l'expérience, nous utiliserons quatre détecteurs avec des saphirs de 250 g ayant des seuils de 0,5 keV et des résolutions de 0,2 keV à 1 keV, ce qui permettra d'obtenir une sensibilité aux masses de WIMPs inférieures à 10 GeV. En 1995 la mise au point des principaux composants cryogéniques a été achevée et testée avec succès. L'installation de notre équipement au LNGS a commencé.

M. Pavan - Recherche des WIMPs au détecteur thermique TeO2 : Premiers résultats de l'expérience Milan/Gran Sasso.
Notre groupe a poursuivi une expérience de désintégration Double Beta en utilisant un ensemble de quatre détecteurs thermiques à 340 g de TeO2. Le spectre de bruit (1548 heures) du bolomètre montrant les meilleures performances (ie 2 keV FWHM de résolution d'énergie et un seuil 13 keV de seuil d'énergie) a été analysé pour trouver des signaux de WIMPS. Devant le manque d'indication d'interactions des WIMPS, nous donnons une limite sur la section efficace indépendante du spin des noyaux Te et O.

Fushimi - Recherche de matière noire spin-couplée au moyen de grands scintillateurs de volume.
On recherche les WIMPs ayant des interactions dépendantes du spin au moyen de scintillateurs NaI et CaF_2de grand volume. Les diffusions inélastiques avec un couplage vecteur axial sur ^{127}I par diffusion inélastique de la matière noire ont été étudiées, ce qui a permis d'obtenir une nouvelle limite sur la présence de la matière noire spin-couplée. Un nouveau système de détection CaF_2 a été mis au point pour étudier la diffusion élastique des WIMPs spin couplées .

T. Ali. Limites de la mesure des interactions dans l'expérience britannique sur la matière sombre.
La détection directe des WIMPs est possible lorsque leur section efficace de diffusion avec des nucléons est comparable à celle d'un neutrino.
Nous rapportons ici une amélioration significative de nos résultats précédents sur le taux des interactions qui dépendent du spin de la matière sombre. Une limite supérieure de 2σ par rapport aux taux de 0,14 événements/Kg.keV.jour est atteinte (95% dc).

L. Bergstrom. Détection indirecte de matière cachée supersymétrique.
Nous passons en revue les différentes méthodes de détection des particules supersymétriques neutres qui pourraient constituer le halo de notre galaxie.

J. Edsjö. Annihilation des WIMPs dans le soleil et dans la terre.
L'annihilation des WIMPs au centre du soleil ou de la terre peut produire des neutrinos énergiques. La distribution angulaire des muons provenant de l'interaction neutrino-nucléon est étudiée à l'aide de simulation Monte Carlo. Nous montrons que des télescopes à neutrinos pourraient permettre une mesure de la masse des WIMPs à partir de la largeur de distribution angulaire des muons.

Y. Sacquin. NESTOR.
L'expérience NESTOR a pour but de construire un détecteur de neutrinos de haute énergie, installé en mer à grande profondeur. Les objectifs de physique couverts par le projet concernent l'astrophysique, la physique des neutrinos, ainsi que la recherche de matière noire sous la forme de neutralinos massifs. Les performances attendues du détecteur ainsi que l'état actuel du projet sont présentées. Les principaux travaux en cours sont passés en revue, ainsi que les étapes essentielles du projet. La collaboration NESTOR entend participer à la réalisation d'un futur détecteur mondial de 1 km^3.

C. Spiering. Etat des projets BAIKAL et AMANDA.
Nous résumons l'état actuel des télescopes de neutrinos en construction au lac Baikal, en Sibérie, et au Pôle Sud. Les résultats de l'expérience Baikal comprennent un premier candidat de neutrino et des limites supérieures de monopôles magnétiques. En 1996, le groupe Amanda a développé un détecteur à une profondeur de 1500 à 2000M. Nous résumons la configuration et les capacités du détecteur.

E. Paré. Les télescopes atmosphériques Cerenkov peuvent-ils résoudre le problème de la matière noire.
En observant l'univers par le biais de son émission de gammas à très haute énergie, la technique du Cerenkov atmosphérique (ACT) peut contribuer à résoudre le problème de la matière noire. Deux possibilités sont étudiées. Les WIMPs peuvent être observés par les gammas engendrés lors de leurs annihilations.

En mesurant comment les gammas venant d'AGN sont absorbés par la lumière émise par les galaxies dans le milieu intergalactique, on peut donner des informations sur l'époque de formation de ces galaxies. L'état de l'art de la technique sera exposé ainsi que les développements et projets en France et ailleurs.

G. Tarlé. Détection indirecte des WIMPs par les positrons et antiprotons cosmiques : situation actuelle et perspectives d'avenir.
Nous passons en revue la situation actuelle de la détection indirecte des WIMPs utilisant les antiprotons et positrons présents dans la radiation cosmique. De très récentes mesures du rapport antiproton/proton à basses énergies (<1 GeV) sont en excellent accord avec de nouveaux calculs de production secondaire sur le milieu interstellaire. Le rapport dépasse les valeurs le splus optimistes prédites pour l'annihilation des WIMPs à toutes énergies, réduisant la probabilité de détection des WIMPs par cette méthode. De récentes mesures de la fractionde positrons n'ont pas oncfirmé l'augmentation intéressante précédemment rapportée à hautes énergies. (> 10 GeV). Il faudra des mesures nettement améliorées, avec de plus amples échatnillons statistiques et un meilleur rejet des hadrons, avant d'attteindre les prédictions le splus optimistes pour l'annihilation des WIMPs en paires d'électron et positron.

R. E. Lanou - HERON, détecteur de matière noire ?
HERON (Helium : Détection de neutrinos à l'aide de rotons) est un projet visant à mettre au point un nouveau détecteur de neutrinos solaires des chaînes p-p et ^{7}Be. Nous présentons ici de nouveaux résultats qui montrent une assymétrie dans la distribution de l'émission des rotons par les particules arrêtées, et une possibilité de détecter simultanément la fluorescence ultra violette. Ces caractéristiques sont potentiellement valables pour la détection des neutrinos solaires et la question se pose, de savoir si la même technique peut être utilisée pour la détection des WIMPs de la matière noire.

W. Forman - La matière noire dans les galaxies les plus anciennes.
Nous décrirons ici les progrès récents dans la connaissance de la matière noire des galaxies de type avancé (elliptique et lenticulaire). Les observations ROSAT de la couronne de rayons X autour de ces galaxies montrent qu'elles possèdent un halo de matière noire ayant un rapport masse/luminosité de $\approx 100 M/L$. Nous comparons ces mesures à celles obtenues pour des systèmes plus grands, groupes et amas. Nous résumons enfin les implications cosmologiques de la matière noire, du gaz chaud et de la répartition de la masse stellaire.

J. Lu. Axions de SUSY SM et SUSY GUT comme candidats à la matière noire.
Récemment l'intérêt pour l'axion a grandi car c'est un candidat à la matière noire froide. Nous discutons des propriétés et des méthodes de détection des axions.

E. Lançon. Résultats récents du LEP 1.5 au CERN
Les expériences au LEP 1.5 se sont déroulées sur 3 semaines à l' automne 1995. Les mesures relevées aux énergies de centre de masse 130 et 136 GeV, avec une luminosité totale intégrée de 5.7 pb^{-1}, ont été analysées par les quatre expériences LEP. `De nouvelles particules et l'indication de nouveaux phénomènes ont été activement recherchées. Aucun événement n'ayant été trouvé, cela a permis de fixer des limites sur les masses et les sections efficaces de production des sleptons et des estops, des charginos et des neutralinos. Les domaines, exclus lors du LEPI, ont été largement étendus. Des recherches de particules lourdes produites par paire et se désintégrant dans le mode hadronique, ont été aussi entreprises dans la topologie des quatre jets. Les résultats préliminaires de chacune des quatre expériences LEP sont présentés séparément.

II. COSMOLOGIE ET FORMATION DES STRUCTURES

J.S. Bagla. Contraintes d'observation sur Ω et H_0.
Nous considérons deux classes de modèles cosmologiques, d'une part les modèles plats avec constante cosmologique, d'autre part les modèles ouverts sans constante cosmologique. Nous utilisons le modèle standard CDM pour contraindre la formation des structures. Comparés aux modèles ouverts, les modèles avec constante cosmologique ont une plus grande chance de survie.

M. Lachièze-Rey. Formation des galaxies : la théorie linéaire.
Je présente les grandes lignes du développement des fluctuations qui ont été à l'origine de la formation des galaxies et des grandes structures cosmiques. Je restreins la présentation aux premier smoments de ce développement, caractérisé par la possibilité d'appliquer une approximation linéaire qui donne son nom à la phase correspondante de la dynamique. Je décris aussi les anisotropies attendues dans le fond diffus cosmologique, qui sont engendrées durant cette phase linéaire. Enfin, j'indique très brièvement une piste possible pour examiner le développement des fluctuations ultérieur à cette phase linéaire.

W. H. Zurek - La distribution de matière autour de galaxies lumineuses.
Nous allons discuter de Ω_R, mesure de la contribution à la densité de matière cosmique à l'intérieur d'une distance R des galaxies lumineuses. Si la matière de l'univers était fortement liée aux galaxies, Ω R croîtrait rapidement jusqu'à la densité moyenne Ω à une échelle comparable aux dimensions d'un halo typique de galaxie. Utilisant des simulations numériques, nous démontrons que , dans les modèles CDM standard et de faible densité, seulement la moitié de la masse totale de l'univers se trouve dans une région d'environ 1 Megaparsec autour des galaxies ayant une luminosité plus grande que L*. Cette conclusion a une importance considérable pour les mesures de masse conventionnelles qui utilisent un modèle de masse ponctuelle pour les galaxies.
Nous explorons deux de ces mesures : la méthode de moindre action et le théorème du viriel cosmique. Dans le premier cas, la méthode ne fonctionnera probablement pas à une échelle comparable à la distance typique entre galaxies, mais devrait fonctionner correctement pour estimer la masse d'un groupe isolé de galaxies ou d'un amas pauvre. Dans le deuxième cas, nous trouvons que la grande masse qui n'appartient pas aux galaxies risque de fausser l'estimation de Ω.

F. Bernardeau. Formation d'une structure à grande échelle dans le régime quasi linéaire.
La compréhension de la formation des grandes structures requiert la résolution d'équations non linéaires couplées décrivant l'évolution des champs de densité et de vitesse cosmologiques. C'est un problème compliqué qui, ces dix dernières années, a été traité essentiellement avec des simulations numériques à N corps. Il y a cependant un régime, le régime dit quasi linéaire, pour lequel les fluctuations relatives de densité sont inférieures à l'unité en moyenne. Il est alors possible d'utiliser des techniques de théorie des perturbations où les développements perturbatifs sont faits par rapport aux fluctuations initiales. Je présente ici les résultats majeurs qui ont été obtenus dans ce régime.

S. Colombi. Le régime hautement non linéaire et les simulations à N corps.
Je discute de la dynamique à grande échelle dans un univers en expansion composé de matière non collisionnelle autogravitante, en prêtant particulièrement attention au régime hautement non linéaire, qui est généralement traité en utilisant des simulations à N corps. Mon approche est statistique, en termes de fonctions de corrélation et de fonction de distribution du champ de densité lissé. Après avoir rappelé les préjugés théoriques attendus (autosimilarité, équilibre statistique local), je discute des mesures dans des simulations à N corps.

G. Tormen - Simulations hydrodynamiques de formation de galaxies.
Ce rapport est une brève introduction à l'hydrodynamique numérique dans un contexte cosmologique, destinée aux non-spécialistes. Nous présentons d'abord les principaux processus de formation des galaxies. Puis nous introduisons les équations des fluides et nous décrivons leur utilisation dans les codes numériques grâce aux méthodes basées sur une grille Eulérienne et sur l'hydrodynamique des particules lissées. A titre d'exemple, nous exposons quelques résultats d'une simulation SPH d'un amas de galaxies.

J. Loveday. SDSS (Sloan Digital Sky Survey) : Etat et perspectives.
Le SDSS a pour objet de cartographier π steradians de l'univers local. Une matrice de détecteurs CCD en mode balayage produira une image digitalisée du ciel avec cinq filtres différents et avec une magnitude limite r'≈23.

Une étude spectroscopique sera faite sur une sélection de 10^6 galaxies et 10^5 quasars. Je décris ici l'état d'avancement du relevé, qui devrait commencer à produire des observations en 1997, et les perspectives qu'il ouvre pour contraindre les modèles de matière noire dans l'univers.

G. Mamon
Les relevés en infrarouge proche DENIS et 2MASS sont présentés avec leurs applications à l'astronomie extragalactique et à la cosmologie. Les perspectives d'un suivi spectroscopique rapide d'un échantillon sélectionné en infrarouge proche sont illustrés avec des simulations Monte Carlo.

N. Mandolesi. Projet spatial COBRAS/SAMBA
La mission COBRAS/SAMBA de l'ESA a pour objet de fournir des cartes étendues et précises, à plusieurs fréquences, des anisotropies de la radiation du fond diffus cosmologique, avec une sensibilité angulaire de 0,1 à 180 degrés. Cela permettra une parfaite identification des perturbations primordiales de densité qui, en s'amplifiant, ont formé les structures à grande échelle observées dans notre univers. Les observations de COBRAS/SAMBA serviront de test décisif pour les théories en compétition décrivant l'origine des structures à grande échelle et permettront une définition très précise de tous les paramètres cosmologiques.

J. Delabrouille. Que peut-on encore apprendre du fond de rayonnement cosmologique ?
Les anisotropies du fond de rayonnement cosmologique sont engendrées par toute une variété de processus physiques, depuis l'Univers Primordial dans la première seconde du Big-Bang, jusqu'à nos jours. Le satellite COBRAS/SAMBA a pour objectif la mesure de ces anisotropies avec une précision inégalée. Cette mesure permettrait de contraindre très fortement les modèles cosmologiques et, dans la plupart des cas de figure, de déterminer les valeurs des principaux paramètres cosmologiques comme H_0, Ω_0 ou Λ, avec une précision de l'ordre du pourcent.

R. Pain. Mesures de Q_0 avec des Supernovae de type Ia.
Dans une première étape de notre programme de recherche de supernovae à grand décalage vers le rouge, nous avons découvert 7 supernovae. Les spectres et courbes de lumière indiquent que pratiquement toutes sont de type Ia avec des décalages vers le rouge variant entre z=0,3 et z=0,5. Ces supernovae peuvent être utilisées comme indicateurs de distance pour mesurer le paramètre cosmologique q0. Nous présentons ici les stratégies d'observation, l'analyse ainsi que les résutats préliminaires concernant la mesure de q_0.

M.Y. Khlopov - Théorie de la formation de structure à grande échelle basée sur principes de la physique.
La théorie moderne de formation de structure à grande échelle contient plusieurs éléments qu'on peut relier aux mécanismes physiques de l'inflation, de la baryosynthèse et à la physique de la matière noire. Ces mécanismes sont prédits dans la théorie des particules mais ne sont pas basés sur l'expérience. Le cadre cosmologique, le spectre des fluctuation primordiales et la fonction de transfert, définis par les propriétés des particules de la matière noire, le rapport de densité matière baryonique/matière noire, et leur distribution relative, ont des motivations et des champs d'application différents et sont généralement traités séparément. Le traitement physique autocohérent correspond à un cadre théorique unique qui offre à la fois une base définie pour les scénarios de structure à grande échelle et des possibilités pour des études détaillées par la combinaison d'effets additionnels, prédits dans leur cadre théorique respectif. Nous proposons en exemple des scénarios cosmologiques, basés sur le principe d'unification horizontale.

III. MESURES QUANTIQUES

W. H. Zurek. Un test précis de décohérence.
Le mouvement d'une particule chargée au dessus d'un conducteur plat est dissipé par la résistance ohmique due aux courants images. Cette interaction entre la particule et le conducteur produit aussi de la décohérence, qu'on peut détecter dans les franges d'interférence formées par des faisceaux de particules qui ont traversé le conducteur.

Puisque la densité de courant dans le conducteur est très sensible à la hauteur de la particule, on peut contrôler le degré de décohérence, et de cette façon, étudier le phénomène de décohérence du domaine quantique jusqu'au domaine classique.

P.L. Knight. Mesures d' états quantiques à l'aide de «règles quantiques».
Cet article décrit comment le bruit quantique contribue aux processus de détection en optique quantique, et comment il est représenté dans l'espace des phases. Les mesures simultanées des variables canoniques sont déterminées non seulement par les propriétés du système, mais également par celles de la « règle quantique » (ou cosystème utilisé pour réaliser la mesure).

F. Ricci. Un oscillateur mécanique macroscopique au niveau de la limite quantique : un schéma opératif.
Avec un système à Evasion de la RétroAction (ERA), il est possible de contourner la limite de sensibilité des détecteurs linéaires de déplacement. Nous présentons ici les prévisions sur les performances d'un système ERA optimisé par rapport au prototype que nous avons déjà mis en fonction en système classique. Nous montrerons que ce système expérimental permet de dépasser la limite quantique standard.

Golub. Nouvelles questions en optique neutronique.
L'optique neutronique au moyen d'instruments variant avec le temps, comme des lames ou des miroirs mobiles, offre de nouvelles possibilités d'étude sur, par exemple, l'équation de Schrödinger appliquée aux particules massives, les transitions entre états classique et quantique pour un système simple et les propriétés de cohérence des faisceaux de neutrons.
Le matériel optique à spin est une sorte d'ordinateur analogique pour effets optiques où le spin moyen des particules classiques arrivant à un point précis peut servir de modèle pour les effets de diffraction et d' interférence en optique standard. Nous introduisons le concept et proposons quelques applications.

A. M. Steane. Correction des erreurs quantiques.
Pour des sytèmes classiques, les effets indésirables du bruit peuvent être contrôlés par une stabilisation active. Dans le cas des systèmes quantiques, pour lesquels une évolution unitaire doit être préservée, une telle stabilisation n'est pas possible. Ceci semble anéantir les espoirs que l'on fonde sur les possibilités d'un ordinateur qui fonctionnerait sur le mode quantique. Cependant, en adaptant des méthodes classiques de correction d'erreur, l'information quantique peut être stabilisée activement. La communication d'états quantiques en présence de bruit est ainsi rendue possible, et il est probable que l'élaboration de calculs au niveau quantique puisse bénéficier de ces nouvelles techniques.

S. Reynaud. Synchronisation d'horloges avec des champs quantiques.
La comparaison de temps par tranfert de signaux lumineux joue un rôle central dans la théorie de la relativité. La nature quantique des champs utilisés dans une procédure de synchronisation d'horloge conduit à rediscuter les notions de temps et de fréquence ainsi que leurs transformations dans les changements de référentiel. On utilise l'invariance dans les transformations conformes de la théorie des champs quantiques de masse nulle, de la définition du vide et du nombre des particules. Ces transformations permettent en particulier de représenter les référentiels uniformément accélérés. On définit les références de temps ou de fréquence qui peuvent être encodées dans un pulse et transférées d'un observateur à un autre. ces références sont des observables quantiques, exprimées en fonction des générateurs associés aux concepts classiques de temps ou de fréquence, et présentent des fluctuations quantiques compatibles avec la quatrième relation de Heisenberg.
Les relations de commutation caractéristiques de l'algèbre conforme permettent ensuite de déduire les transformations de ces observables dans un changement de référentiel. La transformation du temps lors du changement d'un référentiel inertiel vers un référentiel uniformément accéléré diffère de la prédiction classique associée aux propriétés de covariance, tout en respectant la compatibilité des transformations des références de temps et de fréquence.

A. Lambrecht. Radiation de mouvement émise par une cavité oscillante.
Nous avons effectué des calculs permettant d'évaluer de manière quantitative l'influence de la finesse de la cavité. La cavité présente en effet des résonances mécaniques pour lesquelles le rayonnement est multiplié par la finesse, par rapport au cas d'un miroir oscillant dans le vide. Le spectre du rayonnement est lui aussi changé puisqu'il contient maintenant des pics de résonance optique. La cavité rayonne non seulement lorsque sa longueur mécanique est modulée, mais également quand elle oscille globalement, sa longueur restant constante. Ce résultat paradoxal du point de vue de la mécanique - la cavité rayonne alors qu'elle se déplace dans le vide sans référence autre que le vide lui-même - se rencontre déjà pour un miroir unique, mais on bénéficie ici en plus de l'effet d'exaltation résonnante.
Grâce à cet effet d'exaltation, la mise en évidence des effets mécaniques dissociatifs des fluctuations du vide paraît pour la première fois réalisable avec les ordres de grandeur que nous avons obtenus.

Puppo - Mesure de la force de Casimir dans une cavité optique par un schéma de transduction résonnant.
En 1948, Casimir et Polder démontrèrent l'existence de l'effet Casimir comme conséquence des fluctuations du vide du champ électromagnétique. Nous présentons une étude expérimentale pour mesurer dynamiquement la force de Casimir dans une cavité confocale en utilisant un schéma de transduction résonnant.

A. Ekert. Sécurité de l'information et mécanique quantique.
Existe-t-il des codes réellement secrets ? L'ancien art de coder l'information a, par le passé, toujours été égalé par l'ingéniosité de ceux qui brisaient ces codes. Aujourd'hui toutefois, une combinaison de physique quantique et de cryptographie promet d'anéantir, peut-être définitivement, les espoirs des espions potentiels. Nous présentons une brève introduction de cryptographie quantique et commentons les récents développements expérimentaux.

I. Cirac - Etude de la dissipation et des méchanismes quantiques avec des ions piégés.
Nous proposons une méthode pour mettre en équation l'expérience du phénomène d'incohérence dans la mécanique quantique. Elle ouvre un nouveau champ d'expérience pour des phénomènes qui, jusqu'à présent, n'ont été étudiés que sur le plan théorique.

J.M. Raimond & S. Haroche. L'ordinateur quantique : rêve ou cauchemar ?
Un ordinateur quantique est un rêve de théoricien : il calcule infiniment plus vite qu'un ordinateur classique, et peut résoudre rapidement des problèmes difficiles, comme la factorisation des grands nombres. Mais c'est aussi un cauchemar pour les expérimentateurs : la décohérence condamne en fait toute tentative de construire un calculateur assez grand.
Nous discutons ces problèmes à la lumière de quelques expériences récentes en optique quantique.

IV. LA GRAVITATION EXPERIMENTALE

H. Heitmann. VIRGO : un détecteur à laser des ondes de gravitation.
Le détecteur franco-italien d'ondes gravitionnelles VIRGO est en cours de construction. Il s'agit d'un interféromètre laser de type Michelson avec des bras de 3 km de longueur, contenant des cavités Fabry-Pérot pour augmenter la longueur effective des bras, et utilisant la technique de recyclage de lumière pour augmenter la puissance effective du laser. Cet article donne brièvement un résumé des différentes sources de bruit limitant la sensibilité des détecteurs interférométriques d'ondes gravitationnelles, et décrit les principaux éléments constituant l'interféromètre VIRGO.

A. Lazzarini. Le point sur le projet LIGO.
LIGO va contribuer à ouvrir le champ de l'astrophysique des ondes gravitationnelles. LIGO va détecter, en utilisant l'interférométrie optique entre des masses test libres, les minuscules distorsions dynamiques dans l'espace-temps, produites par les ondes gravitationnelles.

J. Frieben. Etoiles à neutrons triaxiales - Une source éventuelle de rayonnement gravitationnel.
Les étoiles à neutrons triaxiales pourraient être des sources de rayonnement gravitationnel importantes pour la prochaine génération de détecteurs interférométriques d'ondes gravitationnelles comme LIGO, VIRGO ou GEO600. Nous étudions l'instabilité séculaire du mode barre des étoiles à neutrons en rotation rapide par l'intermédiaire d'une analyse perturbative de modèles stellaires axisymétriques « exacts » en relativité générale.

R.W.P. Drever. Nouveaux concepts pour la détection d'ondes gravitationnelles par interférométrie laser.
Nous décrivons certains nouveaux concepts permettant d'étendre les performances et la gamme de fonctionnement des interféromètres à laser pour la détection des ondes gravitationnelles. Parmi ceux-ci, il y a notamment l'utilisation de réseaux de diffraction holographique pour le traitement des miroirs, ce qui permettra d'atteindre une plus grande puissance et de réduire le bruit thermique. Nous explorerons aussi les possibilités d'étendre la sensibilité des interféromètres aux ondes gravitationnelles de basse fréquence en utilisant des techniques comme la lévitation magnétique.

P. Fayet. Les motivations théoriques pour des tests du principe d'équivalence et la recherche de nouvelles forces à longue portée.
Nous discutons les raisons théoriques de considérer de nouvelles forces à longue portée, induites par des particules de spin 1 ou 0.
Les couplages d'un nouveau boson de jauge de spin 1, le U, sont obtenus en identifiant les symétries $U(1)$ supplémentaires susceptibles d'être jaugées et en tenant compte des effets de mélange avec le Z. La « charge de cinquième force », qui s'exprime comme
$Q_5 = xB + yL + zQ_{el}$, apparaît dans le cadre de la grande unification comme effectivement proportionnelle au nombre de neutrons. La force correspondante, qui dépend de la composition, s'ajouterait à la gravitation, conduisant à des violations apparentes du Principe d'Equivalence, et en général aussi, de la loi de Newton en $1/r^2$. Son intensité effective $\bar{\alpha}$ est reliée à sa portée λ et à l'échelle de brisure F de la symétrie $U(1)$ supplémentaire par $\bar{\alpha} \approx 1/(\lambda^2 F^2)$.
Un théorème d'équivalence montre qu'un boson U de spin très léger ne se découple pas lorsque sa constante de jauge s'annule, mais se comporte comme un boson pseudoscalaire de masse quasi nulle. Les expériences de physique des particules imposent alors à F d'être supérieur à l'échelle électrofaible. Pour les grandes valeurs de F, nous obtenons le mécanisme du « U invisible ».
Les échanges de U peuvent induire des forces entre spins, et même une interaction « masse-spin » violant la symétrie CP. Ces forces proportionnelles à $1/F^2$, pourraient être plus facilement détectables que dans le cas de l'axion (pour lequel λ et F sont liés par une relation de proportionnalité).

K. Nordtvedt. Réduction des erreurs sytématiques dans les données de poursuite de la lune au laser et de Mars au radar.
La poursuite laser de la lune et la poursuite radar de Mars (dans la mission Viking) ont produit les deux meilleurs ensembles de données pour tester la gravitation relativiste. L'analyse de ces données fait cependant la part belle à de possibles erreurs sytématiques. J'esquisse ici un travail dont le but est de rationaliser le traitement des erreurs systématiques en incluant des représentations paramétrisées de biais plausibles ou attendus. Ceci devrait permettre de ramener les barres d'erreur finales plus près des erreurs formelles associées à l'ajustement par moindres carrés.

E. Fishbach. Masse minimum des neutrinos.
L'échange de neutrinos de masse nulle entre des fermions lourds produit une force à deux corps de longue portée. Nous montrons que la force analogue à plusieurs corps conduit à une grande densité d'énergie dans les étoiles des neutrons. Pour réduire la densité de l'énergie à une valeur acceptble, le neutrino doit avoir une masse minium qui est environ de 0,4 eV/c^2.

R. D. Newman. Perspectives sur les expériences de physique de la gravitation avec des balances de torsion cryogéniques.
Nous discutons des perspectives sur différentes expériences (tests perfectionnés du principe d'équivalence faible, recherches d'anomalies de la gravité Newtonienne (ou nouvelles forces), et mesure de la constante de gravitation G), qui utilisent des balances de torsion fonctionnant à températures cryogéniques. Les basses températures ont les avantages d'un haut Q mécanique avec un bruit thermique faible, une grande stabilité mécanique et thermique.

R. C. Ritter. Recherche de couplage anormal de spin avec la matière noire hypothétique dans notre galaxie.
Un pendule de torsion avec des masses contenant ≈1023 spins électroniques intrinsèques alignés, mais sans interaction magnétique mesurable, a été utilisé dans le passé pour chercher d'hypothétiques forces faibles dépendant du spin. On a reconfiguré cet instrument pour chercher si une interaction de spin à grande distance existe dans une direction privilégiée, comme celle qui pourrait être due à un halo de matière invisible non baryonique. Grâce à la rotation de la terre, la direction dans laquelle le pendule est sensible parcourt le ciel et l'angle qu'elle fait avec la direction de l'angle galactique a une variation périodique. Après 790 jours sidéraux d'observation, l'analyse statistique fixe une limite inférieure pour un tel signal au niveau de 3% de l'accélération gravitationnelle vers le centre galactique. Pour étudier la contribution de possibles effets systématiques, on a appliqué un filtre sidéral qui doit supprimer beaucoup de ces effets de signature diurne. Au moment où nous écrivons ces lignes, la moyenne cumulative, après filtrage d'un signal vers le centre galactique est (0,025 +/- 0,011) fois la valeur de l'attraction gravitationnelle, où l'incertitude rapportée est celle obtenue par le traitement statistique.

M. Moore. Mesures précises de forces faibles.
Après plus de 200 ans, la balance ou pendule de torsion est encore l'un des moyens les plus précis de mesurer de faibles forces à portée macroscopique. Nous avons développé et testé une technique de mesure à l'aide d'un pendule de torsion, dans laquelle les effets dus aux variations de la température de la fibre sont supprimés par plus de quatre ordres de grandeur. Nous observons l'amplitude de la deuxième harmonique du mouvement du pendule. Une mesure récente sur 21 jours en utilisant un pendule nul (pas de signal attendu) avec Q = 4000 a donné une amplitude de 0,4±0,6 nanoradians pour le deuxième harmonique. Ceci équivaut pour un pendule d'équilibre à un déplacement de 1,6±2,4 nanoradians.
Cette nouvelle technique suggère qu'une amélioration significative dans la précision de ces expériences de physique fondamentale peut être atteinte avec un Q relativement élevé et un appareil à une température d'environ 300K.

M. Bantel. Pour une redétermination de G.
Nous discutons une expérience en développement à l'UCI pour déterminer la constante gravitationnelle G.

Un pendule de torsion à T≈4K mesure le décalage de fréquence d'oscillation provoqué par deux masses sources en anneaux fixées à l'extérieur du dewar cryogénique. Les anneaux vont produire un gradient très pur de champ quadrupolaire au niveau du pendule. Le pendule doit être une plaque mince de quartz suspendue dans un plan vertical par une fibre en Al5056 sous vide absolu. Les mesures seront ainsi très peu sensibles à l'incertitude de la répartition de la masse et de la localisation du pendule.
Les mesures seront faites avec le pendule oscillant avec plusieurs grandes amplitudes (2.57, 4.21, 5.81...radians) pour lesquelles le décalage de fréquence atteint un maximum en fonction de l'amplitude. Cela améliore le rapport signal/bruit par rapport à des expériences de petite amplitude, rend les résultats insensibles aux erreurs dans la détermination de l'amplitude, et offre un moyen de vérification cohérent pour un certain nombre d'effets systématiques. L'expérience se déroulera dans un ancien bunker de missiles Nike qui offre un bruit sismique très bas et de faibles variations temporelles des gradients de gravitation.

J.H. Gundlach - Mesure de la constante G de Newton au moyen d'une balance de torsion en rotation.
Nous présentons ici une nouvelle méthode de mesure de la constante G de Newton au moyen d'une balance de torsion en rotation permanente qui est mise en oeuvre avec une rétroaction en vitesse.
Cette méthode permet de réduire la sensibilité aux incertitudes systématiques dominantes des expériences précédentes. Nous avons procédé avec succès à des tests préliminaires qui démontrent la faisabilité de cette nouvelle technique.

V. CONFERENCES D'INTERET GENERAL

E. Fischbach. Dix années de cinquième force.
Nous présentons une revue sur la phénoménologie de la cinquième force et sur les contraintes expérimentales.

L. Celnikier. Cherche source des rayons cosmiques ultra énergétiques.... désespérément.
Depuis quelques années, nous savons que le spectre du rayonnement cosmique s'étend au delà de 10^{20} eV, et nous avons de fortes présomptions qu'il s'agit de protons. Ces particules ne sont associées avec aucune source céleste connue ; d'ailleurs, on ne sait pas par quel processus des protons peuvent être accélérés à de telles énergies. Le projet AUGER vise à mettre l'étude de cette nouvelle fenêtre sur des bases solides.

LIST OF PARTICIPANTS

Tariq ALI
Imperial College
Astrophysics GP, Blackett Lab,
Prince Consort Rd
SW7 2BZ London, U.K.
t.ali@ic.ac.uk

Blas CABRERA
Stanford University
Physics Dpt
CA 94305 Stanford, USA
cabrera@leland.stanford.edu

Reza ANSARI
Université Paris-Sud
LAL, Bât. 200
91405 Orsay Cedex, France
ansari@frcpn11.in2p3.fr

Fabien CAVALIER
Université Paris-Sud
LAL
Bât. 208
91405 Orsay Cedex, France
cavalier@frcpn11.in2p3.fr

Jasjeet BAGLA
I.U.C.A.A.
Post Bag 4
Ganeshkhind
411007 Pune, India
jasjeet@incaa.ernet.in

Ludwik CELNIKIER
Observatoire de Meudon
DAEC
5 Place Janssen
92195 Meudon Principal Cedex, France
celnikier@mesiob.obsm.fr

Michael BANTEL
University of California
Dept of Physics & Astronomy
CA 92717-4575 Irvine, USA
mbantel@uci.edu

Gabriel CHARDIN
CE Saclay
DAPNIA/SPP
91191 Gif sur Yvette Cedex, France
chardin@hep.saclay.cea.fr

Lars BERGSTRÖM
Stockholm University
Dept of Physics
Box 6730
11385 Stockholm, Sweden
LBE@PHYSTO.SE

Vérène CHAZAL
IPN Lyon
Université Cl. Bernard
43 Bd du 11 Novembre 1918
69622 Villeurbanne Cedex, France
chazal@frcpn11.in2p3.fr

Francis BERNARDEAU
CE Saclay
SPT
Orme des Merisiers
91191 Gif sur Yvette Cedex, France
fbernard@amoco.saclay.cea.fr

Ignacio CIRAC
Univ. of Castilla- La Mancha
Depart. de Fisica aplicada
Campus Universitario S/N
13071 Ciudad Real, Spain
icirac@teazer.uibk.ac.at

Luc BLANCHET
Observatoire de Meudon
92195 Meudon Principal Cedex, France
blanchet@obspm.fr

Stephane COLOMBI
University of Toronto
CITA
60 St George St
ON M5S 1A7 Toronto, Canada
colombi@cita.utoronto.ca

Michel BRUNE
ENS
Labo. Kastler Brossel
24 Rue Lhomond
75231 Paris Cedex 05, France

Christian COUTURES
CE Saclay
DAPNIA/SPP
91191 Gif sur Yvette Cedex, France
coutures@dphyxz.saclay.cea.fr

Thibault DAMOUR
IHES
35 Route de Chartres
91440 Bures sur Yvette, France
damour@ihes.fr

Artur EKERT
Oxford University
Clarendon Laboratory
Keble Road
OX1 3PU Oxford, U.K.
ekert@physics.ox.ac.uk

Gavin DAVIES
Imperial College
Physics Dpt, HEP group,
Prince Consort Rd
SW7 2BZ London, U.K.
g.j.davies@ic.ac.uk

John ELLIS
CERN
Theory Division
1211 Genève 23, Switzerland
johne@crnvma.cern.ch

Alvaro DE RUJULA
CERN
Theory Division
1211 Genève 23, Switzerland
derujula@NXTH13.CERN.CH

Orrin FACKLER
Lawrence Livermore Nat. Lab.
7000 East Avenue
CA 94506 Livermore, USA
fackler@llnl.gov

Nathalie DELABROUILLE
CE Saclay
DAPNIA/SPP
91191 Gif sur Yvette Cedex, France
nathalie@hep.saclay.cea.fr

Pierre FAYET
ENS
Labo. de Physique Théorique
24 Rue Lhomond
75231 Paris Cedex 05, France
fayet@physique.ens.fr

Jacques DELABROUILLE
Université Paris-Sud
IAS
Bât.121
91405 Orsay Cedex, France
jacques@ias.fr

Ephraim FISCHBACH
Purdue University
1396, Physics Dept,
IN 47907-1396 West Lafayette, USA
ephraim@physics.purdue.edu

Ronald W.P. DREVER
Caltech
Gravitational Physics
130-33
CA 91125 Pasadena, USA
rdrever@caltech.edu

William FORMAN
Smithonian Astrophysical Obs.
60 Garden Street
MA 02138 Cambridge, USA
wrf@cfa.harvard.edu

Georgi DVALI
CERN
Theory Division
1211 Genève 23, Switzerland
dvali@suzya11.cern.ch

Joachim FRIEBEN
Observatoire de Meudon
DARC
5 Place Janssen
92195 Meudon Principal Cedex, France
frieben@obspm.fr

Joakim EDSJÖ
Uppsala University
Dept of Theoretical Physics
Box 803
751 08 Uppsala, Sweden
edsjo@teorfys.uu.se

Ken-Ichi FUSHIMI
Tokushima University
Dept of Physics
1-1 Minami Josanjima-cho
770 Tokushima, Japan
fushimi@ejsn.phys.sci.osaka-u.ac.jp

Denis GARRETA
CE Saclay
DAPNIA/SPP
91191 Gif sur Yvette Cedex, France
garreta@frcpn11.in2p3.fr

Gilles GERBIER
CE Saclay
DAPNIA/SPP
91191 Gif sur Yvette Cedex, France
gerbier@hep.saclay.cea.fr

Yannick GIRAUD-HERAUD
Collège de France
LPC
11 Place Marcelin Berthelot
75231 Paris Cedex 05, France
giraud-heraud@cdf.in2p3.fr

Robert GOLUB
Hahn-Meitner-Institut
Glienicker Strasse 100
14109 Berlin, Germany
golub@HMI.DE

Paolo GONDOLO
University of Oxford
Dept of Physics, Theoretical Physics,
1 Keble Rd
OX1 3NP Oxford, U.K.
gondolo@thphys.ox.ac.uk

Philippe GRANGIER
Institut d'Optique
BP 147
91403 Orsay Cedex, France
philippe.grangier.@iota.u-psud.fr

Bruno GUIDERDONI
IAP
98 bis Bd Arago
75014 Paris, France
guider@iap.fr

Jens GUNDLACH
University of Washington
Nuclear physics Lab.
Box 354290
WA 98195-4290 Seattle, USA
jens@phys.washington.edu

Serge HAROCHE
E.N.S
Labo. Kastler Brossel
24 Rue Lhomond
75231 Paris Cedex 05, France
haroche@physique.ens.fr

Henrich HEITMANN
Université Paris-Sud
LAL
Bât. 208
91405 Orsay Cedex, France
heitmann@lalcls.in2p3.fr

Yusuf JAFRY
ESA/ESTEC
Space Science Dept
Postbus 299
2200 A9 Noordwijk, Netherlands
yjafry@so.estec.esa.nl

David Ian JONES
University of Wales
Dept of Physics & Astronomy
College of Cardiff, PO Box 913
CF2 3YB Cardiff, U.K.
d.jones@astro.cf.ac.uk

Gerard JUNGMAN
Syracuse University
Dept of Physics, High Energy Th. Group
201 Physics Building, Room 309
NY 13244-1130 Syracuse, USA
jungman@npac.syr.edu

Guinevere KAUFFMANN
MPI für Astrophysik
Karl-Schwarzchild-Str. 1
85748 Garching bei München, Germany
gamk@mpa.garching.mpg.de

Maxim KHLOPOV
Cosmion
Center of Cosmoparticle Physics
Miusskaya Pl. 4
125047 Moscow, Russia
khlopov@khlopov.rc.ac.ru

Peter KNIGHT
Imperial College
Blackett Laboratory
Prince Consort Road
SW7 2BZ London, U.K.
p.knight@ic.ac.uk

Marc LACHIEZE-REY
CE Saclay
DAPNIA/SAp
Orme des Merisiers
91191 Gif sur Yvette Cedex, France
marclr@ariane.saclay.cea.fr

Astrid LAMBRECHT
Ludwig-Maximilian Universität
Schellingstr. 4/III
80799 München, Germany
asl.@ ipp.garching.mpg.de

Eric LANCON
CE Saclay
DAPNIA/SPP
91191 Gif sur Yvette Cedex, France
elancon@hep.saclay.cea.fr

Robert LANOU
Brown University
Dept of Physics
Box 1843
RI 02912 Providence, USA
lanou@brownvm.brown.edu

Albert LAZZARINI
Caltech
102-33 East Bridge Lab.
M/S 51-33
CA 91125 Pasadena, USA
kzz@ligo.caltech.edu

Stefan LOCHMANN
ZARM
AM Fall Turm
28359 Bremen, Germany
lochmann@zarm.uni-bremen.de

Jon LOVEDAY
Fermilab
PO Box 500
IL 60510 Batavia, USA
loveday@fnal.gov

Jizong LU
Shangai Teachers University
Dept of Physics
100 Guilin Rd
200234 Shangai, China
jil@mppmu.mpg.de

Anthony LUN
Monash University
Dept of Mathematics
Wellington Rd
VIC 3168 Clayton, Australia
lun@vaxc.cc.monash.edu.au

Gary MAMON
IAP
98 bis Bd Arago
75014 Paris, France
gam@iap.fr

Nazzareno MANDOLESI
Institute T.E.S.R.E
CNR
Via Gobetti, 101
40129 Bologna, Italy
reno@botes1.tesre.bo.cnr.it

Tahar MELLITI
Université Paris VI
L.G.C.R, Tour 22/12
Boite 142, 4 Place Jussieu
75005 Paris 05, France
tam@ccr.jussieu.fr

Peter MINKOWSKI
University of Bern
Institut für Theoretische Physik
Sidlerstr. 5
CH- 3012 Bern, Switzerland
mink@butp.unibe.ch

John MOFFAT
University of Toronto
Dept of Physics
60 St George Street
ON M5S 1A7 Toronto, Canada
moffat@medb.physics.utoronto.ca

Michael MOORE
University of Washington
Dept of Physics
Box 351560
WA 98195 Seattle, USA
mmoore@phys.washington.edu

Luigi MOSCA
CE Saclay
DAPNIA/SPP
91191 Gif sur Yvette Cedex, France
mosca@dapcls.saclay.cea.fr

Riley NEWMAN
University of California
Dept of Physics & Astronomy
CA 92717-4575 Irvine, USA
rdnewman@uci.edu

Eckehart NOLTE
TU München
Fakultät für Physik
E 15
85747 Garching, Germany
nolte@physik.tu.muenchen.de

Kenneth NORDTVEDT
Lohrmann Observatory
Mommsenstr.13
01062 Dresden, Germany

Keith OLIVE
University of Minnesota
School of Physics & Astronomy
MN 55455 Minneapolis, USA
FVS6308@VX.CIS.UMN.EDU

Reynald PAIN
Université Paris VI et VII
LPNHE
4 Place Jussieu
75252 Paris Cedex 05, France
rpain@lpnax1.in2p3.fr

Eric PARÉ
LPNHE
Ecole Polytechnique
91128 Palaiseau Cedex, France
pare@polhp1.in2p3.fr

Maura PAVAN
INFN
Gruppo Alte Energie
Via Celoria 16
20133 Milano, Italy
pavan@vaxmi.mi.infn.it

Mario PIMENTA
LIP
Av. Elias Garcia 14 -1
1000 Lisboa, Portugal
pimenta@VAXLIP.LIP.PT

Martin PLENIO
Imperial College
Blackett Laboratory
Prince Consort Road
SW7 2BZ London, U.K.
m.plenio@ic.ac.uk

Paola PUPPO
University of Florence
Dept of Physics
Piazzale E. Fermi 2
50125(ARCETRI) Firenze, Italy
puppo@fi.infn.it

Jean-Michel RAIMOND
E.N.S
Labo. Kastler Brossel
24 Rue Lhomond
75231 Paris Cedex 05, France
JMR@physique.ENS.fr

Serge REYNAUD
Labo. Kastler Brossel
UPMC case 74
4 Place Jussieu
75252 Paris Cedex 05 Cedex, France
sr@spectro.jussieu.fr

Fulvio RICCI
University of Rome
Dept of Physics
Piazza A. Moro 2
00185 Rome, Italy
ricc@roma1.infn.it

Rogers RITTER
University of Virginia
Jesse W. Beams Lab. of Physics,
Mc Cormick Rd
VA 22901 Charlottesville, USA
rcr8r@virginia.edu

Barry ROBERTSON
Queen's University
Dept of Physics
SNO, Stirling Hall
ON K7L 3N6 Kingston, Canada
bar@mips2.phy.queensu.ca

Yves SACQUIN
CE Saclay
DAPNIA/SPP
91191 Gif sur Yvette Cedex, France
sacquin@afsmail.cern.ch

Stephan SCHILLER
University of Konstanz
Fakultät für Physik
M696
78434 Konstanz, Germany
stephan.schiller@uni-konstanz.de

Tom SHUIT
U. C. Berkeley
Center for Particle Astrophysics
301 Le Conte
CA 94720 Berkeley, USA
shutt@physics.berkeley.edu

Daniel SNOWDEN-IFFT
Amherst College
Physics Dept
Merrill Science Center
MA 01002 Amherst, USA
ifft@dilbert.physics.amherst.edu

Sofoklis SOTIRIOU
Athens University
Physics Laboratory
104 Solonos Str.
10680 Athens, Greece
sotiriou@grphla.hepl.uoa.ariadne-t.gr

Christian SPIERING
DESY- Zeuthen
Platanenallee 6
15738 Zeuthen, Germany
csspier@ifh.de

Andrew STEANE
Oxford University
Clarendon Laboratory
Parks Rd
OX1 3PU Oxford, U.K.
a.steane@physics.ox.ac.uk

Gregory TARLE
University of Michigan
Dept of Physics, 2071 Randall Lab.
500 E. University
MI 48109-1120 Ann Arbor, USA
tarle@mail.physics.lsa.umich.edu

Giuseppe TORMEN
Max-Planck Institut für Astrophysik
Postfach 1523
85740 Garching, Germany
bepi@MPA.GARCHING.MPG.DE

Jean TRAN THANH VAN
Université Paris-Sud
LPTHE
Bât. 211
91405 Orsay Cedex, France
trantv@frcpn11.in2p3.fr

Ludovic VAN WAERBEKE
Observatoire Midi-Pyrénées
14 Av. Edouard Belin
31400 Toulouse, France
waerbeke@obs-mip.fr

Michael VOGELEY
Space Telescope Science Institute
3700 San Martin Drive
MD 21218 Baltimore, USA
vogeley@stsci.edu

Gregor WEIHS
Innsbrück University
Institut für Experimentalphysik
Technikerstr. 25
6020 Innsbrück, Austria
gregor.weihs@uibk.ac.at

Yoshihisa YAMAMOTO
Stanford University
E.L. Ginzton Laboratory
CA 94305-4085 Stanford, USA
yamamoto@loki.stanford.edu

Anton ZEILINGER
Innsbrück University
Institut für Experimentalphysik
Technikerstr. 25
6020 Innsbrück, Austria
anton.zeilinger@uibk.ac.at

Ludwig ZERLE
MPI für Physik
Werner Heisenberg Institut
Föhringer Ring 6
80805 München, Germany
zerle@vms.mppmu.mpg.de

Wojciech ZUREK
Los Alamos Nat. Lab.
Theoretical Astrophysics, T 6
MS B288
NM 87545 Los Alamos, USA
whz@LANL.GOV