UNIVERSITÉ DE GENÈVE
Département de physique théorique

FACULTÉ DES SCIENCES
Mme Professeur R. DURRER

Cosmic Microwave Background Anisotropies: Beyond Standard Parameters

THÈSE

présentée à la Faculté des sciences de l'Université de Genève
pour obtenir le grade de Docteur ès sciences, mention physique

par

Roberto TROTTA
de
Italie

Thèse N^o 3534

GENÈVE
Atelier de reproduction de la Section de physique
2004

To Elisa

nobody,not even the rain,has such small hands

E. E. Cummings

Acknowledgments

It is a pleasure to thank the people who have contributed to the realization of this work and who have accompanied me along the way. During these three years, I have enjoyed working with and learning from all the members of the Geneva Cosmology Group and of the group of Michele Maggiore. I have had the chance to collaborate with several people, to whom I would like to express my appreciation: Pedro P. Avelino, Rachel Bean, Rebecca Bowen, Ruth Durrer, Steen H. Hansen, Carlos J. A. Martins, Alessandro Melchiorri, Alain Riazuelo, Graça Rocha, Joseph Silk and Pedro T. P. Viana.

Ruth Durrer has been a wonderful supervisor, never counting the hours she spent on my questions, constantly stimulating my interests while allowing me to pursue my research in all freedom. I thank her for her teaching and for her example. The help of Alain Riazuelo was precious during my first year and gave me a swift start into real research work. I am grateful to Alessandro Melchiorri for involving me in many collaborations which constitute a major part of this thesis. I enjoyed having various interesting and promising discussion with Filippo Vernizzi, Martin Kunz and Céline Boehm, which I am sure one day we will be able to finalize. Sam Leach is guilty to have converted me to the Bayesian school during many restless and humorous discussions. I am indebted to Christophe Ringeval and Thierry Baertschiger for their help in solving my various computer problems, and to Andreas Malaspinas for his Sisyphus work of troubleshooting our computer network. I shall join my thanks to the ones of the cosmology community to Anthony Lewis for developing, supporting and making publicly available the CAMB and COSMOMC codes. I would like to express my gratitude to Prof. Silk, Prof. Courvoisier and Dr. Ferreira for accepting to be part of the jury.

The European Network CMBNet and the Schmidheiny Foundation have provided generous support for some of the collaborations I was involved with, in the form of travel grants. I am indebted to Oxford Astrophysics and to Prof. Silk for their kind hospitality in many occasions, and to Profs. Spergel and Kosowsky and to Princeton University for financial support of my visit.

But there is not only science, even to the life of a PhD student. My time in Geneva would not have been the same without the friendship of Timon "Lincoln" Boehm: I wish him all the best on his new path. I have had the pleasure of sharing many refreshing moments with Stefano Foffa, Marj Tonini, Yasmin Friedmann, Simone Lelli, Anna Rissone, Martin Zimmermann, Davide "Dutturi" Lazzati. To my parents, my affectionate thoughts for their encouraging presence. To Elisa, my fiancée, my deepest gratitude for having been at my side in all marvellous and all difficult moments, and for the ones which are still to come.

This work was much improved both in form and contents by the careful reading of Christophe Ringeval and Elisa Cunial (for the french part), Sam Leach and Ruth Durrer for the English part. I thank them for their time and competence. I alone bear the responsibility for any mistake which might still be present.

Contents

List of Figures

List of Tables

Lo duca e io per quel cammino ascoso
intrammo a ritornar nel chiaro mondo
e sanza cura aver d'alcun riposo
salimmo su, el primo e io secondo,
tanto ch'i' vidi de le cose belle
che porta 'l ciel, per un pertugio tondo;
e quindi uscimmo a riveder le stelle.

Dante, La divina commedia
Inferno XXXIV, 133-139.

Résumé

Introduction

Un des piliers les plus importants pour le développement de la cosmologie moderne est le rayonnement diffus cosmologique. Ce dernier est constitué de photons produits en équilibre thermique durant la phase chaude, après l'explosion primordiale (le Big-Bang). Du fait de l'expansion et du refroidissement de l'univers, dès que la temperature est descendue au dessous de ~ 0.25 eV, les électrons se sont combinés avec les protons et les neutrons, formant principalement de l'hydrogène et de l'helium. Ce procès, qui remonte à 300'000 ans seulement après le Big-Bang, est appelé « recombinaison ». Dès lors, les photons du rayonnement de fond diffus ce sont propagés quasiment librement jusqu'à nous, tout en réduisant leur énergie du fait de l'expansion de l'univers.

L'origine cosmologique du rayonnement diffus est confirmée par ses caractéristiques particulières : il est très homogène et isotrope, avec une distribution spectrale de corps noir et une temperature de 2.728 K. Si ce rayonnement intéresse les cosmologues, c'est que l'on y a observé la présence de petites fluctuations de temperature (« anisotropies ») à hauteur d'une partie pour 100'000. Celles-ci seraient la conséquence de perturbations quantiques primordiales qui ont constitué la base du développement des structures à grande échelle de l'univers. En simplifiant, on pourrait affirmer que les anisotropies dans le fond diffus représentent un instantané de l'univers tel qu'il était au moment où les photons se sont découplés des électrons (ce qu'on appelle le « découplage » ou, quand on s'interesse à la dependence spatiale des fluctuations, la « surface de dernière diffusion »). La distribution statistique des fluctuations de temperature que l'on mesure aujourd'hui, contient différentes signatures qui sont en relation, d'une part avec les conditions initiales de l'univers primordial et, d'autre part, avec le contenu de matière-énergie de l'univers actuel. En comparant les modèles théoriques avec les observations, on peut placer des contraintes sur la valeur des paramètres cosmologiques et étudier la physique de l'univers primordial à des énergies qui ne seront jamais produites par les accélérateurs de particules. Cette démarche est possible depuis 1992, quand le satellite COBE mesura pour la première fois les anisotropies dans le rayonnement diffus (Smoot et al., 1992). Aujourd'hui, des données de très haute qualité ont été recueillies par différents groupes, et particulièrement par le satellite WMAP (Bennett et al., 2003). On s'attend à des autres améliorations dans les prochaines années, notamment grâce à l'observation de la polarisation du fond cosmique. Plusieurs observations indépendantes semblent aujourd'hui être expliquées par un univers avec des sections spatiales plates, qui contient environ 5% de baryons, 25% de matière noire non-baryonique, et dont le restant, 70%, est constitué d'énergie sombre, éventuellement sous la forme d'une constante cosmologique. Les conditions initiales

sont en grande partie adiabatiques, même s'il n'est pas possible d'exclure des contributions de type isocourbure. Le spectre des fluctuations est proche de l'invariance d'échelle et l'on sait que l'univers a été réionisé à un redshift compris entre 7 et 20.

Cette thèse comporte trois parties. Dans la première partie, nous introduisons des éléments de cosmologie standard, ainsi que la théorie linéaire des perturbations, dans le cadre d'un formalisme invariant de jauge. Nous dérivons ensuite l'équation de Boltzmann, qui décrit l'évolution des perturbations d'une composante relativiste et nous discutons le terme de collision dû à la diffusion de Thomson entre photons et électrons. La génération et l'évolution d'une composante polarisée sont également présentées. La deuxième partie est une application de ces résultats au rayonnement diffus cosmique. En premier lieu, différents approximations analytiques permettent de comprendre les mécanismes physiques qui génèrent les fluctuations de temperature sur différentes échelles. En deuxième lieu, la dépendance quantitative des paramètres cosmologiques et des conditions initiales est explorée en détail à l'aide de techniques numériques. La troisième partie porte sur la comparaison entre les observations et quelques modèles qui incluent des paramètres non-standards. Après une introduction des techniques statistiques utilisées, nous poursuivons, dans les chapitres 6 et 7 deux lignes de recherche : la première s'occupe de la possibilité de détecter ou contraindre, par le rayonnement de fond, des phénomènes physiques au-delà du modèle standard ; la deuxième porte sur l'influence des conditions initiales sur le rayonnement cosmique, ainsi que l'étude des conséquences phénoménologiques de l'existence de perturbations de type isocourbure. Les résultats originaux de ces recherches ont été publiés dans neuf articles, en collaboration avec divers chercheurs : Pedro P. Avelino, Rachel Bean, Rebecca Bowen, Ruth Durrer, Steen H. Hansen, Carlos J. A. Martins, Alessandro Melchiorri, Alain Riazuelo, Graça Rocha, Joseph Silk et Pedro T. P. Viana.

Le rayonnement cosmologique de fond diffus

On décrit l'univers primordial comme un mélange à quatre composantes, à savoir des photons, des neutrinos de masse nulle, des baryons et de la matière noire froide (CDM). À une temperature $T \gtrsim 0.25$ eV, les photons sont étroitement couplés aux baryons grâce à leur diffusion de Thomson avec des électrons, et forment un plasma. Les fluctuations de densité dans ce plasma augmentent sous l'effet de la gravitation, mais la compression qui en résulte est contrebalancée par la pression de radiation des photons : des oscillations harmoniques apparaissent. Dans un modèle simplifié, comprenant photons et matière noire – mais où l'on néglige l'effet dynamique des baryons – l'équation d'évolution du contraste de temperature des photons dans la jauge Newtonienne, $\Delta T/T \equiv \Theta$, se réduit à celle d'un simple oscillateur harmonique forcé :

$$\ddot{\Theta} + \frac{k^2}{3}\Theta = -\frac{k^2}{3}\Psi\,, \tag{R.1}$$

où le potentiel gravitationel Ψ joue le rôle de force d'entraînement. Puisque le potentiel est approximativement constant pour les temps qui nous intéressent, la solution générale est

$$\Theta = A\cos(k\eta/\sqrt{3}) + B\sin(k\eta/\sqrt{3}) - \Psi\,, \tag{R.2}$$

η étant le temps conforme et k le vecteur d'onde. Les conditions initiales déterminent le mode, sinus ou cosinus, qui est excité. Les conditions dites « adiabatiques » correspondent

à des fluctuations initiales avec la même dépendance spatiale dans toutes les composantes et excitent le cosinus. Les conditions de type « isocourbure » entraînent par contre des perturbations de l'entropie entre l'une des composantes et les photons. Le mode sinus est excité par une perturbation initiale dans l'entropie de la matière noire.

Quand la temperature descend au dessous de 0.25 eV, les électrons libres sont capturés par les protons (recombinaison de l'hydrogène), l'univers devient transparent et les photons peuvent se propager librement le long de géodésiques. Les fluctuations de densité du plasma et les irrégularités dans le potentiel gravitationnel au moment de la dernière diffusion, restent figées sous forme de perturbations dans la temperature des photons du rayonnement cosmique. Dans la limite d'une recombinaison instantanée, le contraste de temperature actuel dans le rayonnement cosmologique diffus (CMB) se compose de trois termes, qui dans l'espace de Fourier ont la forme suivante :

$$\Theta|_{\eta_0} = \Theta|_{\mathrm{rec}} - \imath \mathbf{k}\cdot\mathbf{n} v_b|_{\mathrm{rec}} + \int_{\eta_{\mathrm{rec}}}^{\eta_0} \mathrm{d}\eta (\dot{\Psi} + \dot{\Phi})\,, \tag{R.3}$$

où η est le temps conforme, η_0 sa valeur actuelle et η_{rec} la valeur à l'époque de la recombinaison. Le premier terme représente les fluctuations de temperature sur la surface de la dernière diffusion (« effet Sachs-Wolfe ordinaire »). Le deuxième terme correspond à l'effet Doppler dû au mouvement relatif du plasma par rapport à l'observateur. Le dernier terme résulte de la dépendance temporelle des potentiels gravitationnels Ψ et Φ tout au long du chemin parcouru par les photons (« effet Sachs-Wolfe intégré »).

Le contraste de temperature sur la sphère céleste en direction $\mathbf{n}$ peut être développé en harmoniques sphériques,

$$\Theta = \sum_{\ell=0}^{\infty} \sum_{m=-\ell}^{\ell} a_{\ell m} Y_{\ell m}(\mathbf{n})\,. \tag{R.4}$$

La fonction de correlation entre deux directions $\mathbf{n}$ et $\mathbf{n}'$ telles que $\mathbf{n}\cdot\mathbf{n}' = \cos\vartheta$ s'écrit alors

$$C(\vartheta) = \frac{1}{4\pi} \sum_{\ell} (2\ell+1) C_\ell P_\ell(\mathbf{n}\cdot\mathbf{n}')\,, \tag{R.5}$$

où nous avons introduit le *spectre de puissance* C_ℓ, défini comme

$$\langle a_{\ell m} \cdot a^*_{\ell' m'} \rangle = \delta_{\ell\ell'}\delta_{mm'} C_\ell\,, \tag{R.6}$$

où P_ℓ sont les polynômes de Legendre. Si les fluctuation suivent une distribution Gaussienne, le spectre de puissance contient la totalité de l'information statistique. Le spectre de puissance est dominé par des effets différents sur des échelles angulaires différentes. À grande échelle (c'est-à-dire pour $2 \leq \ell \lesssim 20$), on observe un plateau qui dépend du spectre des fluctuations primordiales. Une contribution de l'effet Sachs-Wolfe peut se rajouter si la constante cosmologique ou la courbure de l'univers sont différentes de zéro. Pour une séparation angulaire $\vartheta \lesssim 1^\circ$ ou $\ell \gtrsim 200$, on retrouve une série de pics qui correspondent aux oscillations de $\Theta|_{\mathrm{rec}}$ dans (R.3), ce qui suit de l'équation (R.2) évaluée au moment de la recombinaison. À une échelle encore plus petite, $\vartheta \lesssim 0.2^\circ$, le spectre s'efface de façon exponentielle à cause du couplage imparfait entre photons et baryons. Cet effet est connu sous le nom d' « atténuation de Silk ».

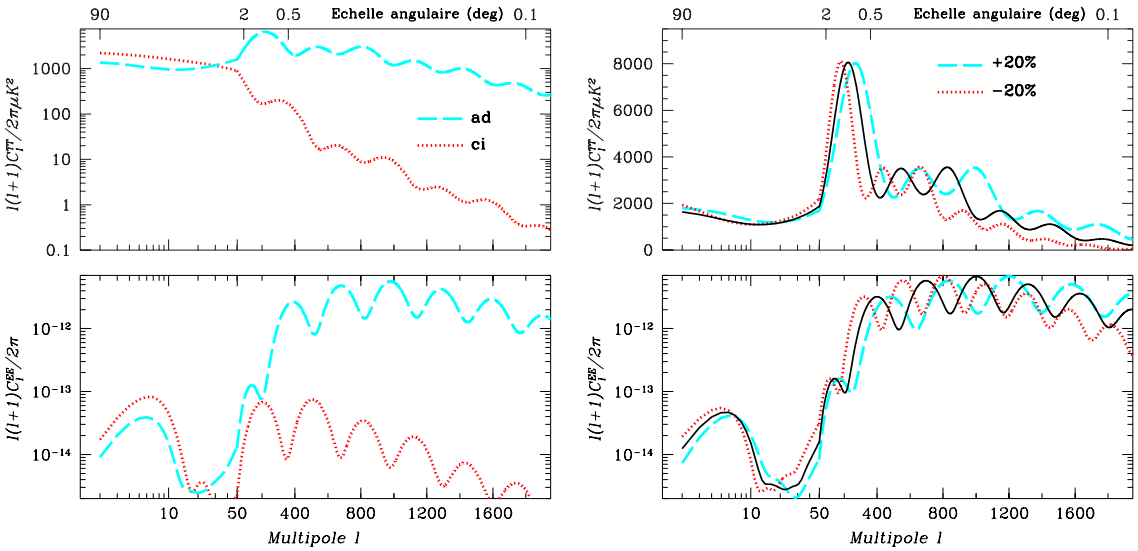

FIG. R.1 – Spectres de puissance de la temperature (en haut) et de la polarisation (en bas). À gauche, pour les même valeurs des paramètres cosmologiques mais avec conditions initiales adiabatiques (ad) ou isocourbure CDM (ci). À droite, pour des conditions initiales adiabatiques mais pour des valeurs différentes de la distance angulaire de la surface de dernière diffusion. Les spectres de puissance sont sensibles aux conditions initiales aussi bien qu'aux paramètres cosmologiques.

La section efficace de la diffusion de Thomson dépend de l'état de polarisation des photons. Une distribution quadrupolaire de la temperature lors de la recombinaison, génère une polarisation linéaire dans le fond diffus d'environ 10% d'amplitude par rapport à la temperature. Le spectre de puissance de la composante polarisée est défini de façon analogue à celui de la temperature et il contient de l'information complémentaire. La polarisation du fond diffus a été récemment observée (Kovac et al., 2002).

La position et la hauteur relative des pics dans les spectres de temperature et de polarisation, sont déterminées par quatre quantités observables et par les conditions initiales. Les quatre quantités sont : la distance angulaire de l'horizon du son évaluée au moment de la recombinaison, l'époque d'égalité entre matière et radiation, la densité baryonique et l'échelle d'atténuation (Hu & Dodelson, 2002). Elles sont aussi appelées « paramètres normaux », car elles ont un effet presque orthogonal sur les spectres (Kosowsky et al., 2002). Les conditions initiales, quant à elles, déterminent la phase des oscillations (sinus ou cosinus) et donc la position du premier pic. En outre, le spectre des fluctuations primordiales contrôle la pente générale du spectre de puissance. La Figure R.1 montre, à gauche, les spectres de temperature et polarisation en fonction d'un différent choix de conditions initiales, et à droite, le spectre adiabatique pour des valeurs différentes de la distance angulaire de la surface de dernière diffusion. À partir des observations, on reconstruit avec précision la valeur des paramètres normaux, si l'on considère les conditions initiales les plus simples, à savoir de type adiabatique. La tâche se complique lorsqu'on essaie de déterminer en même temps les paramètres cosmologiques et les conditions initiales, car d'importantes dégénérescences dans l'espace des conditions initiales sont présentes.

Au-delà des paramètres standards

Le rayonnement diffus cosmologique est un excellent instrument pour tester notre paradigme cosmologique. D'abord parce que sa description est basée sur la théorie de la Relativité Générale et sur ses perturbations linéaires, que nous maîtrisons bien. Deuxièmement, les récentes données sont d'une telle qualité qu'elles nous permettent d'observer des effets très subtils, dont on n'espérait pas voir les conséquences il y a seulement quelques années.

Par la suite, Ω_X est le paramètre de densité d'énergie de la composante X aujourd'hui, exprimé en unités de la densité critique $\rho_{\rm crit} \approx 1.88 \cdot 10^{-29}\, h^2$ g/cm^3. Le paramètre d'Hubble actuel s'écrit $H_0 \equiv 100\, h$ km s^{-1} Mpc^{-1}. La densité d'énergie ρ_X s'exprime alors comme $\rho_X(\eta_0) = \omega_X\, 1.88 \cdot 10^{-29}$ g/cm^3, où nous avons introduit $\omega_X \equiv \Omega_X h^2$ et η_0 est la valeur actuelle du temps conforme.

Le contenu de l'univers en énergies relativistes

Dans le § 6.1 nous nous occupons de l'energie relativiste contenue dans l'univers. Nous décrivons le nombre d'espèces relativistes par le nombre équivalent de familles de neutrinos sans masse, $N_{\rm eff}$ (Bowen et al., 2002). Dans le modèle standard de la physique des particules, $N_{\rm eff} = 3$. Toutefois, il existe plusieurs mécanismes qui pourraient donner $N_{\rm eff} \neq 3$ en cosmologie. En particulier, il est important de remarquer que aussi bien la théorie de la nucléosynthèse que le rayonnement diffus, mesurent l'énergie totale sous la forme de particules relativistes à travers l'effet que celle-ci a sur l'expansion cosmique. De ce fait, la nucléosynthèse et le rayonnement diffus, n'évaluent pas seulement les neutrinos actifs. Ces mesures se réfèrent à deux différents moments de l'histoire cosmique : la nucléosynthèse, à un redshift $z \sim 10^{10}$, et le CMB, à $z \sim 1100$. Il se pourrait donc que le résultat obtenu avec le CMB diffère de celui ressorti de la comparaison entre les prédictions de la nucléosynthèse et les observations des éléments légers. Ce dernier donne $N_{\rm eff} < 3.20$ (avec 95% de confiance) (Burles et al., 1999).

En ce qui concerne le rayonnement diffus, on écrit donc la densité d'énergie en particules relativistes sous la forme

$$\omega_{\rm rel} = 4.13 \cdot 10^{-5}(1 + 0.135 \cdot \Delta N_{\rm CMB})\,. \tag{R.7}$$

On dénote par $\Delta N_{\rm CMB}$ la différence entre le nombre effectif de neutrinos relativistes et le nombre standard, c'est-à-dire $N_{\rm eff} = 3 + \Delta N_{\rm CMB}$. La quantité $\Delta N_{\rm CMB}$ peut être aussi bien positive que négative, ce dernier cas se retrouvant par exemple si les neutrinos ont un potentiel chimique non-nul (Kang & Steigman, 1992). Nous montrons que le paramètre $\omega_{\rm rel}$ est presque dégénéré avec la densité d'énergie en matière ω_m, et cela du moment que le CMB mesure l'époque d'égalité entre matière et radiation, $z_{\rm ég} = \omega_m/\omega_{\rm rel}$, plutôt que les deux termes séparément. En plus, l'utilisation de $\omega_{\rm rel}$ en tant que paramètre non-standard dans l'analyse des données du rayonnement diffus, comporte une dégradation des contraintes sur d'autres paramètres et précisément la courbure de l'univers et l'indice spectral des fluctuations. Nous dérivons ensuite des limites sur $N_{\rm eff}$ à partir d'une compilation des observations du rayonnement diffus effectuées avant WMAP pour obtenir

$$0.04 < N_{\rm eff} < 13.37 \quad (2\sigma)\,. \tag{R.8}$$

Ces limites peuvent être améliorées si l'on inclut des informations supplémentaires, qui découlent d'observations de supernovæ ou de la distribution des structures à grande échelle. En particulier, la limite supérieure devient $N_{\rm eff} < 6.5$ en utilisant les résultats des supernovæ et $N_{\rm eff} < 9.6$ si l'on rajoute au CMB des observations du spectre de puissance de la matière. Même si ces contraintes sont beaucoup moins fortes que celles obtenues par la nucléosynthèse, elles sont néanmoins importantes car totalement indépendantes.

Il est possible de formuler des prédictions détaillées sur les résultats de futures observations avec une technique statistique appelée analyse de la matrice de Fisher. Nous avons formulé de cette façon des prédictions pour le satellite WMAP, qui ont été ensuite confirmées par l'analyse des données. Comme prévu, la meilleure précision des observations sur une échelle plus petite a permis de réduire les dégénérescences, en donnant $\Delta N_{\rm CMB} = 3.4$ (1σ) (Crotty et al., 2003b). Nous avons également appliqué cette méthode au satellite Planck, qui devrait être lancé en 2007 (Planck Website, 2004). La conclusion est que Planck sera capable de déterminer le contenu d'énergie en particules relativistes avec une précision $\Delta N_{\rm eff} \approx 0.24$ (1σ), donc en améliorant les contraintes actuelles d'un facteur 13.

La quantité d'helium primordial et son effet sur le rayonnement cosmique

La nucléosynthèse primordiale prédit l'abondance des éléments légers D, ^{3}He, ^{4}He et ^{7}Li en fonction de deux paramètres seulement, à savoir la densité de baryons et la densité de particules relativistes. Si l'on suit le modèle standard de la physique des particules, avec $N_{\rm eff} = 3$, il s'en suit que la densité de baryons détermine les abondances des quatre éléments légers. Les observations de la présence de ces éléments sont partiellement en désaccord entre elles. On retrouve en particulier des valeurs différentes de la fraction de ^{4}He suivant le groupe de chercheurs qui a effectué l'observation. Si le désaccord entre certaines mesures de la fraction d'hélium et les observations du deuterium était confirmé, cela serait une indication certaine de nouveaux mécanismes physiques à l'oeuvre durant la nucléosynthèse. Cette hypothèse ne peut être vérifiée qu'en disposant d'une mesure indépendante de la fraction d'hélium, car les différences entre les observations astrophysiques pourraient être dues à des erreurs systématiques.

Dans le § 6.2 nous nous intéressons à la possibilité d'utiliser le fond diffus pour mesurer indépendamment la fraction primordiale d'^{4}He, c'est-à-dire la quantité $Y_p = 4n_{\rm He}/n_b$ (Trotta & Hansen, 2004). Ici, $n_{\rm He}$ dénote la densité du nombre d'atomes de ^{4}He et n_b celle de baryons. Une simple estimation montre que l'on s'attend que $Y_p \approx 0.25$. L'effet de Y_p sur les spectres de puissance du CMB se produit à travers des modifications de l'histoire de l'ionisation de l'univers. En particulier, avant la recombinaison de l'hydrogène, la valeur de Y_p contrôlait la densité d'électrons libres dans le régime de couplage fort. Si on suppose que la valeur de Y_p était plus petite, on trouve alors qu'il y avait moins d'électrons libres. De ce fait, le couplage était moins efficace et l'on observe une réduction de la puissance à petite échelle. Cet effet est toutefois très petit, cf. le panneau à gauche de la Figure R.2. L'analyse des données de WMAP, combinées avec d'autres mesures à plus petites échelles, ne donne qu'une détection marginale, à savoir $0.160 < Y_p < 0.501$ (68% de confiance), cf. le panneau à droite de la Figure R.2. Il faut remarquer que les différences entre les mesures astrophysiques de Y_p sont au plus de l'ordre $\Delta Y_p \approx 0.02$, donc le fond diffus est actuellement beaucoup moins précis. Nous estimons toutefois que le satellite Planck pourra réduire l'incertitude à 5%, c'est-à-dire

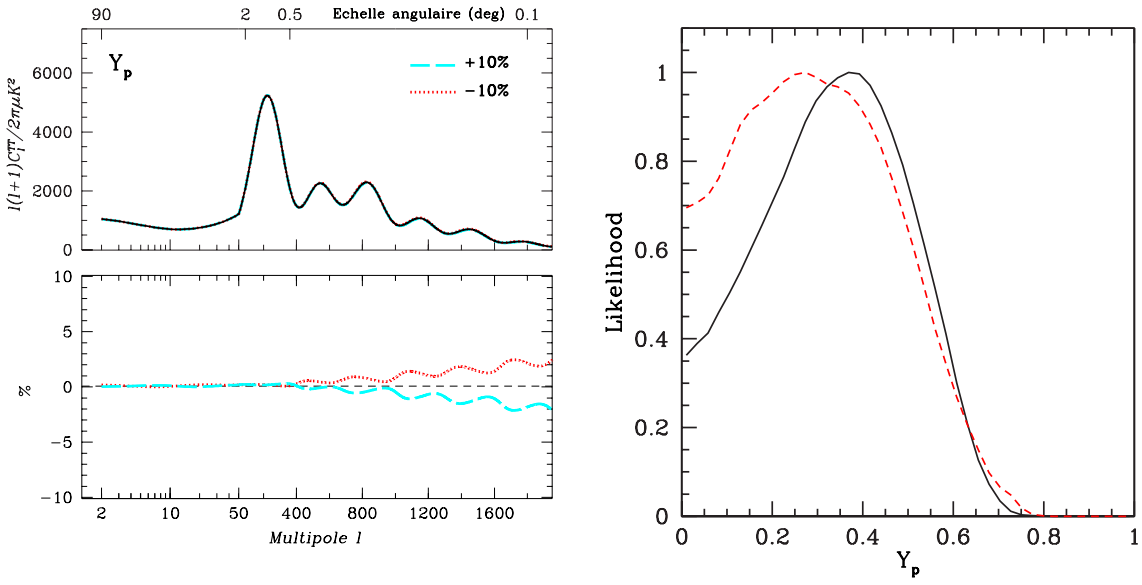

FIG. R.2 – À gauche, on montre l'effet de la fraction d'hélium sur le spectre de puissance de la temperature. À droite, la courbe de la probabilité marginalisée pour la fraction d'hélium Y_p (en noir, ligne continue), obtenue à partir des données sur le fond diffus. L'effet étant très petit, on ne retrouve qu'une détection marginale. La théorie de la nucléosynthèse et les observations d'astrophysique donnent $Y_p \approx 0.24$.

$\Delta Y_p \sim 0.01$, ce qui serait à peine suffisant pour discriminer les résultats astrophysiques qui donnent une petite fraction d'hélium, $Y_p \approx 0.23$ (Olive et al., 1997), et ceux qui donnent une fraction plus importante, $Y_p \approx 0.25$ (Kirkman et al., 2003). Nous montrons que la limite fondamentale à la précision qu'on peut atteindre avec le CMB est de $\Delta Y_p \approx 0.0036$.

Une question importante porte sur les dégénérescences entre Y_p et d'autres paramètres cosmologiques. En particulier, on s'attend à ce que la fraction d'hélium et la densité de baryons soient corrélées entre elles. Il est donc potentiellement important de considérer l'incertitude actuelle sur la valeur de Y_p dans les analyses du fond diffus, afin d'estimer correctement l'erreur sur la densité de baryons. Nous trouvons que cette corrélation n'a pas d'influence sur les données actuelles, mais qu'elle devra être inclue dans l'analyse des données du satellite Planck.

Variations dans le temps de la constante de structure fine

La théorie des supercordes est formulée dans un espace à dix dimensions, qui peut être compactifié et réduit à quatre dimensions. Une des conséquences de cette manipulation est que les constantes fondamentales en dix dimensions acquièrent une dépendance spatiale et temporelle dans notre univers à quatre dimensions. La cosmologie offre la possibilité d'observer de telles variations, et même – en principe – de falsifier la théorie des cordes. Beaucoup d'efforts ont été fournis ces dernières années, pour étudier les variations dans le temps de la constante de structure fine, $\alpha \equiv e^2/(\hbar c)$. Plusieurs méthodes ont été appliquées et le résultat le plus intéressant est probablement celui de Webb et al. (2001) qui ont comparé les spectres de transition atomique de quasars lointains avec ceux mesurés en laboratoire. Ils concluent

que (cf aussi Murphy et al., 2001a,b,c)

$$\frac{\Delta\alpha}{\alpha} = (-0.72 \pm 0.18) \times 10^{-5}\,, \quad z \sim 0.5 - 3.5\,. \tag{R.9}$$

Ceci montre qu'avec plus de 4σ de confiance, α était plus petite dans le passé. Une autre indication vient de l'analyse du réacteur naturel d'Oklo. Dans ce cas, il est toutefois difficile de discriminer entre deux différents solutions, qui mènent l'une à un résultat nul, l'autre à une détection d'une valeur de α qui à $z \sim 0.1$ était plus grande qu'aujourd'hui (Fujii, 2002). Dans beaucoup de modèles, la variation temporelle de α est déterminé par un champ scalaire, et dans ce cas $\alpha(t)$ doit être une fonction monotoniquement croissante dans le temps (Damour & Nordtvedt, 1993). L'observation d'une constante de structure fine qui était plus grande dans le passé pourrait donc falsifier ces modèles, et serait l'indication d'un nouveau mécanisme physique inconnu à l'heure actuelle.

Dans ce cas aussi, le rayonnement diffus s'offre come possibilité indépendante de tester ces résultats, en mesurant la valeur de α à l'époque de la recombinaison, $\alpha_{\rm rec} = \alpha(z \sim 1100)$. Nous présentons dans le § 6.3 les dernier résultats d'une ample collaboration que nous avons menée avec des chercheurs de Paris, Porto, Oxford, Princeton et Rome (Martins et al., 2002, 2004; Rocha et al., 2004). Nous expliquons comment la valeur de $\alpha_{\rm rec}$ influence l'histoire de l'ionisation de l'univers, et donc le spectre de puissance du CMB. Les deux effets principaux sont le déplacement de l'époque de la recombinaison et le changement de l'épaisseur de la surface de dernière diffusion. Ces modifications sont une conséquence du fait que l'énergie de liaison de l'hydrogène est proportionnelle à α^2. Le spectre de la temperature présente donc une réduction de la puissance à petite échelle et un déplacement des pics vers les petits multipoles si $\alpha_{\rm rec}/\alpha_0 < 1$, où α_0 dénote la valeur de la constante de structure fine aujourd'hui. Ces effets sont montrés dans le panneau à gauche de la Figure R.3. La valeur de $\alpha_{\rm rec}$ est partiellement dégénérée avec la distance angulaire de la surface de dernière diffusion, et l'indice spectral $n_{\rm s}$.

En n'utilisant que les données de WMAP, nous obtenons les contraintes

$$0.94 < \alpha_{\rm rec}/\alpha_0 < 1.01 \quad (95\% \text{ de confiance}). \tag{R.10}$$

Bien que la précision du CMB soit beaucoup moins grande que celle atteinte avec les lignes de transition des quasars, il est important de remarquer que le CMB est sensible à des époques très antérieures, $z \sim 1100$. On peut s'attendre à ce que des variations éventuelles de α soient plus importantes à ce moment-là que pour les époques observées grâce aux spectres de quasars, c'est-à-dire $z \sim 2 - 3$. Nous nous intéressons aussi au rôle de la réionisation. Nous montrons que l'existence d'une époque de réionisation précoce, telle que l'indiquent les données de WMAP, est utile pour augmenter la précision sur $\alpha_{\rm rec}$, si l'on suppose que la réionisation se complète sur une courte échelle de temps (réionisation soudaine).

Nous fournissons des prédictions détaillées concernant la précision des futures observations pour tous les paramètres cosmiques standards et α. Par sa nature, le spectre de polarisation peut donner de meilleures contraintes que celui de la temperature, et donc des observations particulièrement pointues de la polarisation seront très utiles. Le satellite Planck doublera la précision de WMAP pour la majorité des paramètres et réduira ainsi l'erreur sur $\alpha_{\rm rec}$ d'un facteur cinq par rapport à aujourd'hui, en obtenant $\Delta\alpha_{\rm rec} \sim 0.3\%$. Si Planck rejoindra pratiquement la limite de la variance cosmique pour la temperature, il n'en sera pas de

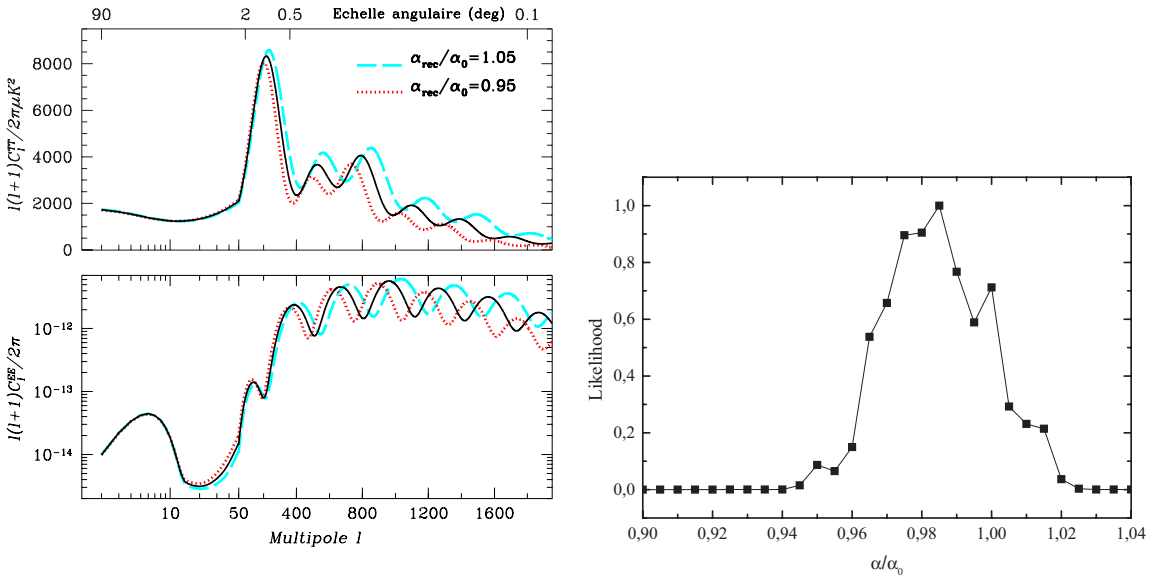

FIG. R.3 – À gauche, effet d'une variation temporelle de la valeur de la constante de structure fine à l'époque de le recombinaison, $\alpha_{\rm rec}$. En haut le spectre de puissance de la temperature, en bas celui de la polarisation. On remarque une modification de la suppression à petite échelle et un déplacement horizontal de la structure des pics. À droite, les contraintes observationelles sur $\alpha_{\rm rec}/\alpha_0$ obtenues des données de WMAP.

même pour la polarisation. En effet, une éventuelle mission ayant une incertitude concernant la polarisation près de la variance cosmique, pourrait encore améliorer la performance de Planck d'un facteur deux ou trois sur la plupart des paramètres, y compris $\alpha_{\rm rec}$.

Tests du paradigme adiabatique

La position du premier pic des oscillations acoustiques – qui se trouve à $\ell = 220.1 \pm 0.8$ (Page et al., 2003) – indique que les conditions initiales sont dominées par le mode adiabatique. Déjà les données de BOOMERanG et MAXIMA avaient permis d'exclure des fluctuations purement isocourbure CDM (Enqvist et al., 2000). Cela ne signifie pas pour autant que des contributions isocourbure ne soient pas présentes. En effet, il n'y pas de raison fondamentale pour croire que l'univers primordial soit décrit par un seul degré de liberté, à savoir un champ inflationnaire unique. Il existe plusieurs scénarios possibles, où par example la présence d'un deuxième champ inflationnaire produit une mixture corrélée de perturbations adiabatiques et isocourbure (Langlois, 1999). Le mécanisme du curvaton (Lyth & Wands, 2002; Moroi & Takahashi, 2001; Enqvist & Sloth, 2002; Lyth et al., 2003), est également à même de générer un mode adiabatique dominant, mais il le fait avec une contribution plus ou moins accentuée d'un mode isocourbure (Lyth et al., 2003). D'un point de vue phénoménologique, on peut donc se demander quelles sont les conditions initiales les plus générales et quelles sont les perspectives de les déterminer en utilisant le fond diffus.

Pour un univers qui contient des photons, des neutrinos sans masse, de la matière noire froide et des baryons, on retrouve cinq solutions aux équations d'Einstein, qui donnent des fluctuations croissantes dans le temps (Bucher et al., 2000). Comme nous le montrons dans

les § 3.2 et § 3.3, ces cinq modes correspondent aux conditions initiales adiabatiques, isocourbure CDM, isocourbure baryons, entropie neutrino et vitesse neutrino. Les modes isocourbure baryons et entropie neutrino sont analogues au mode isocourbure CDM, sauf que la perturbation d'entropie concerne les baryons et les neutrinos, respectivement. Le mode vitesse neutrino est caractérisé par l'absence de fluctuations dans les contrastes de densité des composantes et par une perturbation relative du flux de neutrinos par rapport aux photons. Dans la littérature, on appelle collectivement « modes isocourbure » tous les modes non-adiabatiques, même si tous n'ont pas une perturbation de la courbure nulle au début de leur évolution. Le mode isocourbure baryons étant égal au mode isocourbure CDM, à une constante multiplicative près, on ne peut considérer que quatre modes. Le spectre de puissance le plus général est alors donné par la superposition de ces quatre modes et de leurs corrélateurs. Même si l'on suppose que tous les modes ont le même indice spectral, la description des conditions initiales générales requiert désormais dix paramètres, à savoir les amplitudes des quatre modes, ainsi que celles des six corrélateurs possibles.

Si on admet la possibilité de la présence de modes isocourbures, deux questions se posent naturellement. La première concerne les contraintes qu'il est possible de placer sur les contributions non-adiabatiques. La deuxième porte sur l'estimation des paramètres cosmologiques, indépendamment des conditions initiales. Pour répondre à cette deuxième question, on analyse les données du fond diffus, en marginalisant les conditions initiales générales. De cette façon, on peut vérifier l'importance de la supposition selon laquelle les fluctuations seraient, *a priori*, purement adiabatiques, ce que l'on appelle « le paradigme adiabatique ».

Nous examinons dans le § 7.2 le cas de la densité baryonique et du paramètre de Hubble. Nous trouvons qu'il n'est quasiment plus possible de mesurer ces deux paramètres avec les données antérieures à WMAP si l'on rajoute les modes isocourbure aux conditions initiales (cf. Figure R.4, panneau de gauche). En même temps, il est très difficile de placer des contraintes sur les contributions non-adiabatiques (Figure R.4, panneau de droite). Les conditions initiales générales, ont été introduites dans l'analyse des données du rayonnement de fond pour la première fois dans Trotta et al. (2001). Des travaux plus récents (Crotty et al., 2003a; Bucher et al., 2004) parviennent à des conclusions qualitativement similaires. Ceci montre que la supposition de l'adiabaticité des conditions initiales joue un rôle fondamental dans l'estimation des paramètres cosmologiques. On ne peut donc pas encore parler de « cosmologie de précision » en dehors d'un modèle spécifique pour la réalisation des conditions initiales.

En vue de ce résultat, il est intéressant de se demander si la présence de modes isocourbures a une influence sur la détermination de la constante cosmologique dans les données du rayonnement diffus. La présence d'une constante cosmologique Λ reste un des plus grands mystères de la cosmologie contemporaine. Sa nature démeure inconnue et il n'y a pas d'explication satisfaisante au motif pour lequel de nombreuses observations indiquent $\Omega_\Lambda \sim 0.7$, alors que l'échelle de brisure de la super-symétrie donnerait une estimation de l'ordre $\Omega_\Lambda \sim 10^{58}$. Le fond diffus à lui seul ne peut pas déterminer Ω_Λ à cause d'une dégénérescence géométrique (Efstathiou & Bond, 1999). Comme nous l'avons déjà expliqué, la position du premier pic indique toutefois que $\Omega_\Lambda + \Omega_m \sim 1$. La combinaison de ce résultat avec les observations des supernovæ (Perlmutter et al., 1999; Tonry et al., 2003), des structures à grandes échelles (Tegmark et al., 2004b) et/ou des lentilles gravitationelles (Contaldi

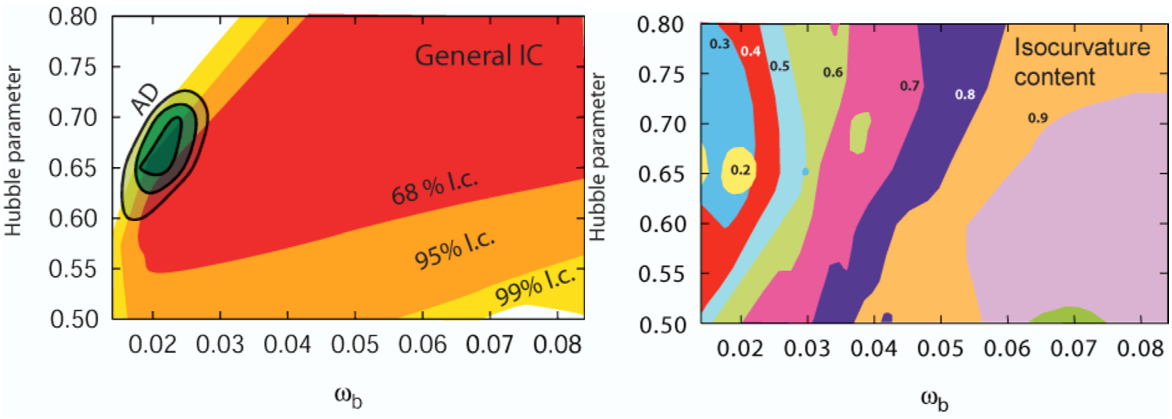

FIG. R.4 – À gauche : régions de probabilité pour la densité baryonique ω_b et le paramètre de Hubble h en supposant des conditions initiales purement adiabatiques (petits contours), et en marginalisant les conditions initiales générales (larges contours). La supposition de l'adiabaticité est essentielle pour obtenir une mesure précise des paramètres cosmologiques. À droite : contribution de modes isocourbures aux modèles qui ajustent mieux les données, où une valeur proche de 0 indique la pure adiabaticité.

et al., 2003) permet de retrouver $\Omega_\Lambda \sim 0.7, \Omega_m \sim 0.3$. Nous avons mis à l'épreuve cette conclusion de deux façon différentes dans Trotta et al. (2003). Ceci est explicité dans le § 7.3, où nous avons considéré des conditions initiales générales et utilisé une approche statistique basée sur la notion de fréquence. Cette dernière donne des intervalles de confiance plus rigoureux par rapport à l'approche Bayesienne, qui est normalement utilisée. Dans notre analyse admettant les modes isocourbures, nous utilisons aussi le spectre de puissance de la matière, qui est comparé aux données du catalogue 2dF Galaxy Redshift Survey. Nous arrivons à la conclusion que même en présence de modes isocourbure, les données du CMB avant WMAP combinées avec 2dF sont difficilement explicables avec un univers du type Einstein–de Sitter, c'est-à-dire $\Omega_\Lambda = 0, \Omega_m = 1$. En particulier, la statistique fréquentiste n'exclue $\Omega_\Lambda = 0$ qu'au niveau de 2σ, mais cela seulement dans le cas où l'on accepte une petite valeur pour le paramètre d'Hubble, à savoir de l'ordre de $h \sim 0.5$. Or, cela serait décidément en contradiction avec la valeur trouvée par le Hubble Space Telescope, c'est-à-dire $h = 0.72 \pm 0.08$ (Freedman et al., 2001).

Si – comme nous l'avons montré – aujourd'hui il n'est pas encore possible d'estimer avec précision les paramètres cosmologiques sans formuler de fortes suppositions sur la nature des conditions initiales, il n'en sera pas de même avec les futures observations du rayonnement diffus. En particulier, la polarisation du fond diffus est d'une importance cruciale pour mieux déterminer les conditions initiales (Bucher et al., 2001). Dans le § 7.4 nous discutons la précision que le satellite WMAP obtiendra sur les quantités directement mesurables avec le CMB, après quatre ans d'observations. Nous comparons les prédictions, en supposant des conditions initiales adiabatiques, au cas où les paramètres sont estimés indépendamment des conditions initiales. Dans ce dernier cas, nous trouvons que WMAP aura une précision de l'orde $10 - 30\%$ (1σ) sur la plupart des paramètres. Nous nous intéressons aussi au satellite Planck, qui aura une résolution bien meilleure pour le spectre de polarisation. Cela lui permettra d'obtenir des erreurs sur les paramètres de l'ordre de $1 - 2\%$, sans aucune hypothèse sur la nature des conditions initiales (Trotta & Durrer, 2004).

Overview and conclusions

Towards a cosmological standard model

The study of cosmic microwave background anisotropies is one of the pillars of modern cosmology. The cosmic microwave background (hereafter CMB) consists of photons left over by the hot phase after the Big-Bang and is very homogeneous and isotropic. Its existence was predicted by Gamov (1946), and accidentally discovered only much later by Penzias and Wilson (Penzias & Wilson, 1965), but it was only in 1992 that the COBE satellite (Smoot et al., 1992) detected the presence of tiny temperature fluctuations (1 part in 100'000), which are thought to have been generated by quantum fluctuations in the very early universe. The observational study of these temperature fluctuations, known as anisotropies, has been a great technological achievement. Over the last ten years, there has been a spectacular advancement in the accuracy of measurements, using ground-based, balloon-born and orbital instruments. The WMAP satellite (Bennett et al., 2003) has recently measured the anisotropies with a precision which, on certain scales, is close to a fundamental statistical limit, called "cosmic variance".

The importance of such a wealth of data for theoretical cosmology cannot be overstated. In a few seconds on a desktop computer, it is nowadays possible to produce accurate numerical predictions of the statistical distribution of the anisotropies on the sky for any cosmological model of interest, i.e. of the CMB angular power spectrum. If the primordial fluctuations are Gaussian distributed, then the power spectrum encodes all of the statistical information: its computation is based on linear perturbation theory and the underlying physics is well understood. The detailed shape of the power spectrum carries characteristic signatures depending on the value of the late Universe cosmological parameters and on the initial conditions for the perturbations. By "late Universe cosmological parameters" we mean the quantities controlling the expansion history of the Universe, i.e. its matter budget, complemented by some description of the reionization history. In the former category, an incomplete list would include the Hubble parameter, the energy density in baryons, cold dark matter and dark energy, the dark energy equation of state parameter (possibly including a description of its time evolution), the neutrino masses and the number of massless families plus the density parameters and effective equation of state of any other exotic form of matter one might wish to include; specifying how the Universe was reionized in the context of stellar evolution theory might require three or four additional parameters, which however usually reduce to the optical depth to reionization or equivalently to the redshift of reionization, as far as the CMB is concerned. Specifying the initial conditions requires the value of "primordial parameters" for the amplitudes of the primordial fluctuations in each of the matter

components and their scale dependence.

The fact that CMB anisotropies are sensitive both to the late Universe cosmological parameters and to primordial parameters means that CMB observations only constrain a (degenerate) combination of both: until now, disentangling the former required rather strong assumptions about the nature of initial conditions. Some guidance is offered by the inflationary paradigm: in its simplest incarnation, the decay of the inflaton field produces adiabatic initial conditions, in which there is no fluctuation in the relative number density of the species, hence no entropy perturbations ("adiabatic"). The presence of entropy fluctuations can excite up to four other non-decaying modes for the perturbations. Those are collectively termed "isocurvature", because in three cases the total matter density is unperturbed and hence there is no curvature perturbation in the spatial sections either. The observation of the first acoustic peak in the CMB power spectrum (Page et al., 2003) at $\ell = 220.1 \pm 0.8$ has substantially confirmed the predominance of the adiabatic mode. However, a subdominant isocurvature contribution to the prevalent adiabatic mode cannot be excluded: after all, there is no compelling reason why the physics of the early universe should boil down to only one degree of freedom.

Even though in principle the number of late Universe parameters can be very large, easily exceeding a dozen, only an handful of them seems to be required by the currently available observational evidence (Spergel et al., 2003; Tegmark et al., 2004b; Liddle, 2004):

- the Hubble parameter $h \sim 0.7$;
- the density parameter for baryons $\Omega_b \sim 0.05$;
- the density parameter for cold dark matter (CDM) $\Omega_{\rm cdm} \sim 0.25$;
- the density parameter for a cosmological constant $\Omega_\Lambda \sim 0.7$;
- the optical depth to reionization $\tau_{\rm re} \sim 0.15$.

Summed together, $\Omega_{\rm cdm} + \Omega_b + \Omega_\Lambda \sim 1$ imply a flat Universe. The crucial point is that for the CMB these results only hold once we make the rather strong assumption of purely adiabatic initial conditions. In that case, the primordial parameters reduce to the spectral index for the fluctuations, $n_{\rm s} \sim 1$, and an overall adiabatic amplitude $A_{\rm AD}$. These two quantities together with the above five late Universe parameters are what we call "standard CMB parameters", because they build the basis of the "concordance model" of present-day cosmology[1].

By combining CMB data with other cosmological and astrophysical measurements – such as galaxy distribution statistics, supernovæ luminosity distance measurements, gravitational lensing statistics, Lyman α absorption lines, local determination of the Hubble parameter, light elements abundance – we have reached an unprecedented precision in determining the standard cosmological parameters, which are now known with an accuracy of a few percent. This is even more astonishing if we think that only ten years ago it was only possible for most parameters to estimate their order of magnitude. Most importantly, various independent

[1] We do not discuss the possibility of gravitational waves, which are indeed predicted by any inflationary scenario; presently there are merely upper limits to their contribution, which could be small enough to be very difficult to detect in the CMB. Our discussion here and in the following focuses on the scalar sector only.

observations – which probe very different epochs of the cosmic history and are based on totally different physical processes – seem to be converging to the same answer.

We are now in a position where we can move on from parameter fitting to model testing: in other words, in order to establish a "cosmological standard model" we need to assess the consistency and completeness of our theoretical framework. In order to be sure that we can trust the error-bars on the standard parameters beyond the quoted statistical error, we have to confront ourselves with the question of possible systematic errors in the measurements on one side, and of hidden flaws in our theoretical interpretation of the data on the other. Given the intrinsic difficulty of many cosmological observations, an assessment of systematic errors for a certain data-set can come from the combination with other, independent measurements of the same quantity. Discrepancies in the results will indicate a flaw in the underlying theory, or in the data, or in both. This is one of the reasons why the comparison of many data-sets is so important, the other being that often the combined data have a superior constraining power due to the breaking of degenerate directions in parameter space. From the point of view of model-building, it is now becoming possible to relax some assumptions which were before necessary in order to extract from the data any information at all, and thereby check whether our results are robust or else whether they critically depend on our prejudices. If it is found that our conclusions depend strongly on the underlying model assumptions, then we need to critically review our theoretical paradigm and open our mind to alternative explicative models.

Testing the concordance model with the CMB

The CMB is an excellent testing ground to carry out this program: our theoretical understanding is based on General Relativity and linear perturbation theory, which suffices to describe almost all of the relevant physical processes. This makes us confident that we understand quite well CMB anisotropies, and we can exploit them to go beyond the standard cosmological parameters in two different ways: the first path leads directly to the primordial Universe, via the dependence of the CMB on the nature of initial conditions; the second approach makes use of the high quality of recent CMB data to look for effects which were previously ignored because thought to be irrelevant, but which are now within the constraining power of the observations. In both cases, the microwave background plays the role of a Universe-sized laboratory for the study of fundamental physics which is often unaccessible to any particle physics laboratory. This work pursues both those aspects, as we detail in the following.

In the first part, we introduce in Chapter 1 the homogeneous and isotropic Friedmann-Robertson-Walker universe, which is the background on which perturbation theory is built, and we briefly present a few other observations which we later compare and combine with the CMB. We then give the derivation of all the relevant perturbation equations needed to describe the CMB in Chapter 2. Those are applied to the temperature fluctuations in the cosmic photons in the second part: in Chapter 3 we obtain under various approximations analytical expressions for the growth of perturbations in an Universe containing photons, cold dark matter, massless neutrinos, baryons and a cosmological constant; in Chapter 4 we present a thorough account of the main features of the CMB temperature and polarization

angular power spectra. In particular, we are concerned with characteristic signatures on the angular power spectra of the standard cosmological parameters, which constitute the basis for their determination using CMB data. We also introduce the most general type of initial conditions, which consist of one adiabatic and four isocurvature modes. The third part focuses on the interplay between theoretical modelling and observational data. The comparison of theoretical models with actual data needs some basis in probability theory and statistics, which we give in Chapter 5, emphasizing their application to the problem of parameter estimation from CMB observations. The last two chapters contain most of the original research work, which is developed along the two lines sketched above: Chapter 6 deals with the observational consequences and constraints when we add to the standard cosmological parameters new quantities describing possible departures from known physics, while Chapter 7 explores the consequences of relaxing the fundamental assumption of adiabaticity.

In § 6.1 we focus on the effective number of massless neutrino families, $N_{\rm eff}$ (Bowen et al., 2002). Although in the standard model of particle physics $N_{\rm eff} = 3$, there are several mechanism which would give $N_{\rm eff} \neq 3$ as measured by the two cosmological probes we discuss, namely Big-Bang Nucleosynthesis (BBN) combined with observations of the light elements abundances, and CMB. This is because both of them are sensitive not only to the number of weakly interacting neutrinos, but rather to the total energy density of relativistic particles which sets the expansion rate at early times, and therefore can constrain e.g. the existence of sterile neutrinos unobservable in Z-decay experiments. Using pre-WMAP CMB data alone, we obtain fairly broad bounds on $N_{\rm eff}$, $0.04 < N_{\rm eff} < 13.37$ with 2σ likelihood content, which are reduced by including prior information coming from supernovæ luminosity distance measurements and large scale structure observations. We show that $N_{\rm eff}$, or equivalently $\omega_{\rm rel} \equiv \Omega_{\rm rel} h^2$, the energy density parameter in relativistic particles, is nearly degenerate with the amount of energy in matter, $\omega_m \equiv \Omega_m h^2$, and that its inclusion in CMB parameter estimation also affects the constraints on other parameters such as the curvature or the scalar spectral index of primordial fluctuations. However, even though this degeneracy has the effect of limiting the accuracy of parameter estimation from the WMAP satellite, we find that it can be broken by measurements on smaller scales such as those provided by the Planck satellite mission. We forecast that Planck will be able to constrain $N_{\rm eff}$ within 0.24 (1σ).

The primordial ^{4}He mass fraction, Y_p, is predicted by BBN along with the abundances of the other light elements as a function of two free parameters, namely the baryon density ω_b and the relativistic energy density $\omega_{\rm rel}$. If we fix $N_{\rm eff} = 3$ and thereby $\omega_{\rm rel}$ as motivated by the particle physics standard model, then in standard BBN the abundances of D, ^{3}He, ^{4}He and ^{7}Li depend on the baryon density alone: comparison with the observed values in astrophysical systems indicates a slight discrepancy, which however presently cannot clearly be ascribed to systematical errors or to deviations from the standard BBN scenario. We explore in § 6.2 the potentiality of using the CMB as a totally independent way of measuring Y_p via its impact on the reionization history, thereby possibly allowing to discriminate between the various hypothesis (Trotta & Hansen, 2004). We find that WMAP data give only a marginal detection, $0.160 < Y_p < 0.501$ at 68% likelihood content. We estimate that the Planck satellite will determine the helium mass fraction within 5% (or $\Delta Y_p \sim 0.01$), which however will only allow a marginal discrimination between different astrophysical measurements. Equally

important, we identify degeneracies between Y_p and other cosmological parameters, most notably the baryon abundance, the redshift and optical depth of reionization and the spectral index; we conclude that even though present-day CMB data accuracy does not require the inclusion of Y_p as a free parameter, the uncertainty of the helium fraction will have to be taken into account in order to correctly estimate the errors on the baryon density from Planck.

The search for observational evidence for time or space variations of the "fundamental" constants that can be measured in our four-dimensional world is an extremely exciting area of current research, with several independent claims of detections in different contexts emerging in the last few years, together with other improved constraints. Most efforts have been concentrating on the fine-structure constant, α, both due to its obviously fundamental role and to the availability of a series of independent methods of measurement. Of particular interest is the result of Webb and collaborators, who claim a 4σ detection of a fine-structure constant that was smaller in the past (Murphy et al., 2003; Webb et al., 2003). Noteworthy among the possibilities of independently check those results is the CMB, which probes $\alpha_{\rm dec}$, the value of α at decoupling, $z \sim 1100$ (Martins et al., 2002, 2004; Rocha et al., 2004). As we show in § 6.3, by analyzing the first year WMAP data for time-variations of α we obtain the constrain $0.95 < \alpha_{\rm dec}/\alpha_0 < 1.02$ with 95% likelihood content, where α_0 denotes the present value. We clarify the issue of degeneracies between α and other standard parameters, and give exhaustive forecasts of the expected performance of the full four year WMAP data, of the Planck satellite and of an ideal CMB experiment. We emphasize the role of polarization measurements to lift flat directions (i.e., degeneracies) in parameter space, and discuss the role of reionization in the determination of $\alpha_{\rm dec}$.

In Chapter 7 we relax the assumption of adiabaticity by allowing for the most general initial conditions (Bucher et al., 2000) and we investigate two complementary aspects: the first is the degradation in the accuracy of the late Universe standard parameters as a consequence of the introduction of new degrees of freedom in the primordial Universe (Trotta et al., 2001); the second is the robustness of the measurement of a non-zero cosmological constant, $\Omega_\Lambda \neq 0$, when different statistical approaches (frequentist rather then Bayesian) are applied to the data, or when general isocurvature modes are included in the analysis (Trotta et al., 2003). We also explicitly test the paradigm of adiabaticity by using CMB observations to put constraints on the isocurvature contribution.

For the first point, the results in § 7.2 demonstrate that the determination of the Hubble parameter and the baryon density from pre-WMAP CMB data is essentially impossible without strong assumptions about the nature of initial conditions. Conversely, it becomes very difficult to put limits on the type of the initial conditions without using external, non-CMB priors on the late Universe parameters. Indeed, the CMB is perhaps the most effective way to directly probe the very early Universe, and thereby constrain or falsify the models for the generation of perturbations. It is therefore very important to extract the most information about the conditions in the early Universe. Adding polarization information greatly enhances the power of the CMB to simultaneously constrain the late Universe parameters and the primordial ones: we show in § 7.4 that the full four year WMAP data will measure orthogonal combinations of the late Universe parameters with an accuracy of the order $10\% - 30\%$ for most parameters even in the general initial conditions case. The

Planck mission will have a better polarization resolution and will be able to do precision cosmology almost independently on the type of initial conditions (Trotta & Durrer, 2004). As for the possibility of mitigating the cosmological constant problem by introducing isocurvature modes, our findings in § 7.3 indicate that $\Omega_\Lambda \neq 0$, as obtained from a combination of CMB and large scale structure data, is indeed robust even in the presence of isocurvature contributions. The more conservative frequentist statistics – as compared to the usual Bayesian approach – excludes $\Omega_\Lambda = 0$ only at the 2σ confidence level for pre-WMAP CMB data combined with the 2dF Galaxy Redshift Survey, but this only if we admit a rather low value for the Hubble constant, $h \sim 0.5$, which would be in contradiction with the result of the Hubble Space Telescope, $h = 0.72 \pm 0.08$ (Freedman et al., 2001).

Outlook and conclusion

The CMB has become a well established tool for the study of our Universe, and an unavoidable testing ground for any theoretical model. The ever improving quality of the data permits on one side to look for new physics in the early Universe, as shown in our study of time variations of α, on the presence of extra relativistic particles and on the existence of non-adiabatic modes; on the other hand, it also requires an upgrade of our modelling, so to properly treat subtle effects such as the uncertainty coming from our unprecise knowledge of the primordial Helium fraction, or from our ignorance on the correct model for the generation of fluctuations. For this reasons, it is important to look ahead, to the goals for the next generation of experiments, and to their potential to constrain or falsify the theoretical models.

More than ever, the central issue is becoming how to efficiently and reliably extract the most information from upcoming high-quality data: there are about 2000 observable independent multipoles for each of the three angular power spectra, namely temperature, E-polarization and temperature-polarization cross-correlation, which however are highly redundant due to the smooth oscillatory nature of the spectra. The amount of information which can be extracted is much less, and can be condensed in maybe a dozen of well-chosen parameters. The best choice for those quantities is the one which takes into account the physics and selects orthogonal directions in parameters space on the basis of fundamental degeneracies. This idea has been a *leitmotiv* of the works presented here, and there is probably still space to apply it further, especially in connection with the primordial parameters.

Despite this encouraging picture, there are still open challenges for our understanding of the Universe: the nature of dark energy and dark matter, the details of the initial conditions and the epoch of reionization, for example. The CMB will provide key advancements on all these issues over the next years. The polarization of the anisotropies has been detected by the experiments DASI (Kovac et al., 2002) and WMAP and will be precisely mapped by the forthcoming experiments PolarBear, Bicep, SPOrt, AMiBA and QUEST, opening up a new line of research and allowing to reconstruct the cosmological parameters with still higher precision. This process will culminate with the European Space Agency satellite Planck (Planck Website, 2004), which starting in 2007 will observe the temperature spectrum with the ultimate possible precision and provide accurate mapping of the polarization as well. In view of this wealth of data, and in order to fully exploit its potential, it is of fundamental

importance that theoretical research on the subject advances accordingly. There is a need of more powerful and efficient computational and statistical techniques which can handle the considerably larger amount of data expected. Also, our theoretical understanding of model-building has to be refined and in particular we need to further develop the interdisciplinary link between models coming from high energy physics, string theory, astrophysics and their observational signature on the CMB. This approach will strengthen the role of the CMB as a universe-size laboratory for investigating the most elusive domains of fundamental physics.

All men, Socrates, who have any degree of right feeling, at the beginning of every enterprise, whether small or great, always call upon God. And we, too, who are going to discourse of the nature of the universe, how created or how existing without creation, if we be not altogether out of our wits, must invoke the aid of Gods and Goddesses and pray that our words may be acceptable to them and consistent with themselves.

Plato
Timaeus

Part I

BASICS

Chapter 1

Introduction

1.1 Notation and conventions

We begin by introducing the notation and conventions which are used throughout this work.

- The metric signature is $-+++$.
- The spacetime metric is denoted by $g_{\mu\nu}$, where the spacetime coordinate are $x^\mu, \mu = 0, 1, 2, 3$. Greek indexes always run from 0 to 3.
- The 3-space of constant curvature has metric γ_{ij}. Latin indexes always run from 1 to 3.
- When we discuss perturbations, the background, unperturbed quantities are denoted by an overline. Therefore for instance $\rho = \bar{\rho} + \delta\rho$, where $\bar{\rho}$ denotes the background energy density and ρ the perturbed (background plus linear perturbation) energy density.
- The overdot " $\dot{}$ " denotes the derivative with respect to conformal time, η.
- Bold character denote the $i = 1, 2, 3$ components of the corresponding 4-vector.
- Unless otherwise stated we use natural units, in which the speed of light, the Boltzmann constant and the Planck constant are unity, $c = k_B = \hbar = 1$.
- The Hubble parameter today is written as $H_0 \equiv 100\, h$ km s^{-1} Mpc^{-1}.
- The symbol Ω_X denotes the density parameter in the component X (where X can stand for baryons, photons, cold dark matter, etc.), expressed in units of the critical energy density. In general, $\Omega_X = \Omega_X(\eta)$, but whenever we omit the explicit time dependence, it is understood that the quantity is evaluated today, i.e. $\Omega_X \equiv \Omega_X(\eta_0)$, where η_0 is the present value of conformal time.
- The critical energy density today is $\rho_{\rm crit}(\eta_0) \approx 1.88 \cdot 10^{-29}\, h^2$ g/cm^3, and the present energy density of component X is written $\rho_X(\eta_0) = \omega_X\, 1.88 \cdot 10^{-29}$ g/cm^3, where we have defined $\omega_X \equiv \Omega_X(\eta_0) h^2$.

1.2 Friedmann-Robertson-Walker cosmology

In this section, we briefly review the standard treatment of an homogeneous and isotropic universe. We present the background Einstein and conservation equations for perfect fluids, along with the unperturbed Boltzmann equation describing relativistic particles.

1.2.1 Einstein equations

The cosmic microwave background is homogeneous and isotropic to better than one part in 100'000. This justifies the assumption that the universe, on large enough scale, can be treated as being homogeneous and isotropic. We then consider a 4-dimensional manifold $\mathcal{M}$ endowed with a metric $g_{\mu\nu}$, so that constant-time hypersurfaces are constant-curvature, maximally symmetric 3-spaces. The Friedmann-Robertson-Walker (FRW) metric reads

$$g_{\mu\nu}\mathrm{d}x^\mu \mathrm{d}x^\nu = -\mathrm{d}t^2 + a(t)\gamma_{ij}\mathrm{d}x^i\mathrm{d}x^j\,, \tag{1.1}$$

with the 3-space metric of curvature $\mathcal{K} = \{0, +1, -1\}$ given by

$$\gamma_{ij}\mathrm{d}x^i\mathrm{d}x^j = \mathrm{d}r^2 + \chi^2(r)(\mathrm{d}\theta^2 + \sin(\theta)^2\mathrm{d}\phi^2)\,. \tag{1.2}$$

Here the *scale factor* $a(t)$ depends only on time, and

$$\chi(r) = \begin{cases} r & \text{for } \mathcal{K} = 0 \text{ (flat universe)} \\ \sin(r) & \text{for } \mathcal{K} = +1 \text{ (closed universe)} \\ \sinh(r) & \text{for } \mathcal{K} = -1 \text{ (open universe)} \end{cases}\,. \tag{1.3}$$

We will mostly work in *conformal time* η, defined through $\mathrm{d}\eta \equiv a^{-1}(t)\mathrm{d}t$, so that the FRW metric reads

$$g_{\mu\nu}\mathrm{d}x^\mu \mathrm{d}x^\nu = a(\eta)(-\mathrm{d}\eta^2 + \gamma_{ij}\mathrm{d}x^i\mathrm{d}x^j)\,. \tag{1.4}$$

Following the assumptions of homogeneity and isotropy, the background energy-momentum tensor, $T_{\mu\nu}$ is bound to be of the perfect fluid form

$$T_{\mu\nu} = (\rho + P)u_\mu u_\nu + P g_{\mu\nu}\,, \tag{1.5}$$

where ρ, P are functions of the conformal time η only, and represent the fluid energy density and pressure, respectively. The fluid 4-velocity is the timelike 4-vector $\mathbf{u}$, with

$$u^\mu = \left(\frac{1}{a}, 0, 0, 0\right) \qquad \text{and} \qquad u_\mu u^\mu = -1\,. \tag{1.6}$$

We suppose that the equation of state of the fluid is of the form

$$P = w(\rho)\rho\,, \tag{1.7}$$

where the enthalpy $w(\rho)$ depends only on the local energy density. In many cases of interest, the enthalpy is simply a constant, in which case it is termed *equation of state parameter*: for cold, non-relativistic, pressureless matter $w_m = 0$ (dust), for relativistic particles $w_r = 1/3$ (radiation) and $w_\Lambda = -1$ for a cosmological constant (vacuum energy). The energy density

of a cosmological constant is contained in $T_{\mu\nu}$, and is of the form $\rho_\Lambda = \Lambda/(8\pi G)$. Another relevant quantity is the *adiabatic sound speed* of the fluid, defined as

$$c_s^2 \equiv \dot{P}/\dot{\rho}\,. \tag{1.8}$$

The Einstein equations

$$G_{\mu\nu} = 8\pi G T_{\mu\nu} \tag{1.9}$$

with the FRW metric (1.4) and the energy-momentum tensor (1.5) yield the two *Friedmann equations*. The first Friedmann equation is a first order differential equation for the *conformal Hubble parameter* $\mathcal{H}(\eta) \equiv \dot{a}/a$

$$\dot{\mathcal{H}} = -\frac{4\pi G}{3} a^2(\rho + 3P)\,. \tag{1.10}$$

The second one is a constraint equation,

$$\mathcal{H}^2 + \mathcal{K} = \frac{8\pi G}{3} a^2 \rho\,. \tag{1.11}$$

An evolution equation for the fluid energy density follows from the 0 component of the energy-momentum conservation equation, $\nabla_\mu T^{\mu\nu} = 0$:

$$\dot{\rho} + 3\mathcal{H}(\rho + P) = 0\,, \tag{1.12}$$

supplemented with the fluid equation of state, Eq. (1.7). If the universe contains (or is dominated by) only one fluid with $w = \text{const}$, it follows from Eq. (1.12) that its energy density behaves as

$$\rho \propto a^{-3(1+w)}\,, \tag{1.13}$$

hence from Eq. (1.10) the scale factor of a flat universe ($\mathcal{K} = 0$) is

$$a = \left|\frac{2A}{1+3w}\eta\right|^{\frac{2}{1+3w}} \quad \text{for } w \neq -1/3\,. \tag{1.14}$$

with $A^2 = 8\pi G/3\rho a^{3(1+w)} = \text{const}$. In particular, in the radiation dominated universe ($w = 1/3$) we have $a \propto \eta$, while in the matter dominated universe ($w \approx 0$) $a \propto \eta^2$.

In the standard cosmological picture, the universe contains non-relativistic, pressureless matter (baryons and cold dark matter), photons, massless neutrinos and a vacuum energy component. In this case, the stress-energy tensor is the sum of the fluid components

$$T^{\mu\nu} = \sum_\alpha T_\alpha^{\mu\nu}\,. \tag{1.15}$$

The Friedmann equations (1.10, 1.11) apply to the total energy density and pressure, which are just the sum of the contributions from each fluid. The energy conservation equation, Eq. (1.12), still applies to the total variables, while in general for each component we have

$$\nabla_\mu T_\alpha^{\mu\nu} = Q_\alpha^\nu\,, \tag{1.16}$$

where the 4 vector $Q_\alpha^{\mu\nu}$ describe the energy-momentum transfer from the component α. The conservation of total energy requires

$$\sum_\alpha Q_\alpha^\nu = 0\,. \tag{1.17}$$

In the general case, the Friedmann equations have to be solved numerically. However, we can easily write down solutions of simple cases. From Eq. (1.13) it follows that for radiation $\rho_r \propto a^{-4}$ while for matter $\rho_m \propto a^{-3}$. Physically, the energy density of matter is diluted by the growth of the physical volume of the 3-space, while for radiation an extra a^{-1} factor comes in from the redshifting of the particles energy. Hence, since a is growing, at early enough time the universe is radiation dominated. The *equality time* is defined as the time at which the two contributions are equal, i.e. $\rho_r = \rho_m$, after which the universe becomes matter dominated. Therefore

$$\frac{a_{\rm eq}}{a_0} = \left.\frac{\rho_r}{\rho_m}\right|_{\eta_0} \approx 3 \cdot 10^{-3}\,, \tag{1.18}$$

or in terms of the *redshift* $z \equiv a_0/a - 1$ we have

$$z_{\rm eq} \approx 3000\,. \tag{1.19}$$

The subscript 0 indicates that the quantity is evaluated today. The numerical estimate comes from the measurement of the present day radiation density in the cosmic microwave background, which together with the assumption of three massless neutrino families yields

$$\rho_r = 7.94 \cdot 10^{-34} \left(\frac{T_{\rm CMB}}{2.737\ {\rm K}}\right)^4 \quad {\rm g/cm^3}\,. \tag{1.20}$$

The matter content of the Universe is obtained from the combination of CMB, large scale structure and supernovæ type IA measurements. We shall see in § 4.2 that the CMB itself is a good probe to determine the redshift of equality.

Since for a cosmological constant $w_\Lambda = -1$, $\rho_\Lambda = {\rm const}$, its contribution is negligible in the early universe, and indeed for a redshift

$$z \gg \left(\frac{\Omega_m}{\Omega_\Lambda}\right)^3 - 1 \approx 0.5\,. \tag{1.21}$$

However, if $\Lambda \neq 0$, the late universe will be dominated by the vacuum energy term. In that case, $a(t) \propto \exp\left[(\Lambda/3)^{1/2} t\right]$ and the expansion becomes exponential (in physical time).

It is customary to introduce the *critical energy density* as the energy density for which the universe is flat

$$\rho_{\rm crit} \equiv \frac{3\mathcal{H}^2}{8\pi G a^2}\,. \tag{1.22}$$

We also define the *Hubble parameter* $H_0 \equiv \mathcal{H}/a_0$ and the *fudge factor* h

$$H_0 \equiv 100\,h\ {\rm km\ s^{-1}\ Mpc^{-1}}\,. \tag{1.23}$$

The critical energy density today then evaluates to

$$\rho_{\rm crit}(\eta_0) \approx 1.88 \cdot 10^{-29}\,h^2\ {\rm g/cm^3}\,. \tag{1.24}$$

At all times, the *density parameters* Ω_X give the contribution of the component X in units

of the critical energy density:

$$\Omega_r(\eta) \equiv \frac{\rho_r}{\rho_{\rm crit}}\,, \tag{1.25}$$

$$\Omega_m(\eta) \equiv \frac{\rho_m}{\rho_{\rm crit}}\,, \tag{1.26}$$

$$\Omega_\Lambda(\eta) \equiv \frac{\rho_\Lambda}{\rho_{\rm crit}} = \frac{\Lambda}{8\pi G\rho_{\rm crit}}\,, \tag{1.27}$$

$$\Omega_\mathcal{K}(\eta) \equiv \frac{-3\mathcal{K}}{8\pi G a^2\rho_{\rm crit}}\,. \tag{1.28}$$

By definition the sum of the density parameters has to be unity

$$\Omega_r(\eta) + \Omega_m(\eta) + \Omega_\Lambda(\eta) + \Omega_\mathcal{K}(\eta) = 1\,. \tag{1.29}$$

The physical energy density of the component X is then given by

$$\rho_X(\eta) = \Omega_X(\eta)\rho_{\rm crit}(\eta)\,, \tag{1.30}$$

and in particular when evaluating this quantity at the present time we define $\omega_X \equiv \Omega_X(\eta_0)h^2$ and write

$$\rho_X(\eta_0) = \omega_X\, 1.88\cdot 10^{-29}\ {\rm g/cm}^3\,. \tag{1.31}$$

The definition (1.28) expresses the energy density due to the curvature of the spatial sections for $\mathcal{K} = \pm 1$. Since $\Omega_\mathcal{K} \propto \mathcal{H}^{-2} \propto \eta^2$, the curvature is always negligible in the early universe. Various cosmological observations indicate that today $\Omega_\mathcal{K} \approx 0$. However, if the universe is not exactly flat, this would imply that at Planck time $|\Omega_\mathcal{K}| \approx \mathcal{O}(10^{-60})$. The smallness of this number is the essence of the "flatness problem". The inflationary mechanism indeed naturally provides a solution for this fine tuning problem: as the universe inflates quasi-exponentially, its curvature is driven to 0.

A key quantity is the *angular diameter distance* $D_A(z)$: consider an object of physical length d sitting at a redshift z_1 (corresponding to conformal time η_1 and radial distance r_1), which is observed at our present position ($z_0 = 0, r_0 = 0$) under an angle θ. Then the angular diameter distance is defined as

$$D_A(\eta_1) \equiv \frac{d}{\theta} = a(\eta_1)\chi(\eta_0 - \eta_1)\,, \tag{1.32}$$

where in the second equality we have used $d = \lambda a(\eta_1)$, with λ the comoving length of the object, and $\theta = \lambda/\chi(r_1)$, noting that $r_1 = \eta_0 - \eta_1$ since light travels on null geodesics. We can now integrate Eq. (1.11) to find

$$\Delta\eta \equiv \eta_0 - \eta_1 = \frac{1}{H_0 a_0^2}\int_{a_1}^{a_0} \frac{{\rm d}a}{\left[\Omega_r + \Omega_m\dfrac{a}{a_0} + \Omega_\mathcal{K}\dfrac{a^2}{a_0^2} + \Omega_\Lambda\dfrac{a^4}{a_0^4}\right]^{1/2}}\,, \tag{1.33}$$

This equation is more conveniently written in redshift space

$$\Delta\eta = \frac{1}{H_0 a_0}\int_0^{z_1} \frac{{\rm d}z}{\left[\Omega_r(1+z)^4 + \Omega_m(1+z)^3 + \Omega_\mathcal{K}(1+z)^2 + \Omega_\Lambda\right]^{1/2}}\,. \tag{1.34}$$

Recall that the quantities Ω_X above are evaluated at the present time. So if we know the physical length of an object at a given redshift, and we measure the angle subtended by

it on the sky, we are in principle able to extract the value of the cosmological parameters using Eq. (1.34). The CMB provides exactly such a standard rod on the sky: the acoustic oscillations of the photon fluid just before recombination have a characteristic length scale, which shows up as the first peak in the angular power spectrum, see § 4.1.2. The redshift of recombination is also known with good accuracy, hence the CMB measures with high precision the angular diameter distance to the last scattering surface. This piece of information alone is however insufficient to reconstruct completely the matter-energy content of the Universe: this problem is known as *geometrical degeneracy*, and it is explained in § 4.1.2.

1.2.2 Boltzmann equation

At early time, the energy density of the universe is dominated by the relativistic species, and to leading order we can neglect in the contribution of non-relativistic components to the total energy. As long as photons are in local thermodynamical equilibrium, the photon temperature T is related to the energy density of radiation by

$$\rho_r = \frac{\pi^2}{30} g_\star T^4 \,, \tag{1.35}$$

where $g_\star$ counts the total number of relativistic degrees of freedom

$$g_\star \equiv \sum_b g_b \frac{T_b^4}{T^4} + \sum_f g_f \frac{T_f^4}{T^4} \tag{1.36}$$

and b and f run over the bosonic and fermionic species respectively. The factors T_b and T_f take into account possible temperature differences between the photons and the other relativistic particles. From Eq. (1.35) and $\rho_r \propto a^{-4}$ it follows that while the photons are in thermodynamical equilibrium, $T \propto 1/a$.

For $T > 4000\text{K} \approx 0.4\text{eV}$ hydrogen nuclei are ionized, and photons are coupled to baryons via non-relativistic Thomson scattering off free electrons, see § 2.2.5. As the temperature drops below 0.30eV, corresponding to $z_{\rm dec} \approx 1100$, almost all the hydrogen nuclei quickly recombine, the mean free path of photons becomes larger than the Hubble length $1/\mathcal{H}$: the universe becomes transparent. This event is called *last scattering* or *decoupling*.

After recombination, the photon distribution function

$$f(\eta, E) = \frac{1}{\exp(E/T) - 1} \tag{1.37}$$

evolves according to the *collisionless Boltzmann equation*, which can be derived by requiring that the total derivative of f with respect to the affine parameter λ vanishes

$$\frac{\mathrm{d}f}{\mathrm{d}\lambda} = 0 \,. \tag{1.38}$$

In general $f = f(\eta, x^i, E, n^i)$, where the momentum 4-vector $p^\mu = (p^0, \mathbf{p})$ is written as

$$p^\mu = \frac{E}{a}(1, \mathbf{n}) \,, \tag{1.39}$$

with

$$p^i = \frac{|\mathbf{p}|}{a} n^i \,, \qquad p^0 = \frac{E}{a} = \frac{|\mathbf{p}|}{a} \,, \tag{1.40}$$

$$\sqrt{p_i p^i} \equiv |\mathbf{p}| \,, \quad n^i n^j \gamma_{ij} = 1 \,. \tag{1.41}$$

From Eq. (1.38) we have

$$\frac{\partial f}{\partial \eta} + \frac{\partial f}{\partial x^i} n^i + \frac{\partial f}{\partial E} \dot{E} + \frac{\partial f}{\partial n^i} \dot{n}^i = 0 \,. \tag{1.42}$$

Because of isotropy, $\partial f / \partial n^i = 0$, while homogeneity implies $\partial f / \partial x^i = 0$. Using the 0 component of the geodesics equation

$$\frac{\mathrm{d}p^\alpha}{\mathrm{d}\lambda} + \Gamma^\alpha_{\mu\nu} p^\mu p^\nu = 0 \,, \tag{1.43}$$

which in the FRW universe reads

$$\dot{E} + \mathcal{H} E = 0 \tag{1.44}$$

we obtain from Eq. (1.42) the background Boltzmann equation

$$\frac{\partial f}{\partial \eta} - \mathcal{H} E \frac{\partial f}{\partial E} = 0 \,. \tag{1.45}$$

This equation is satisfied by any f of the form $f = f(aE)$. We conclude that after decoupling the energy of the cosmic photons is redshifted by the expansion as $E \propto a^{-1}$. The black body distribution, Eq. (1.37), retains its spectrum. The spectrum of the cosmic microwave background photons has been measured very accurately by the FIRAS spectrometer onboard the COBE satellite (Fixsen et al., 1996), and was found to be exceedingly close to thermal. Deviations from a perfect black body spectrum can be measured by the Comptonization parameter y, the chemical potential μ and the parameter Y_{ff} describing contamination by free-free emission. The 95% confidence limits on those parameters are

$$|\mu| < 9 \cdot 10^{-5} \,, \quad |y| < 1.2 \cdot 10^{-5} \,, \quad |Y_{ff}| < 1.9 \cdot 10^{-5} \,. \tag{1.46}$$

After decoupling, T is no longer a temperature in the thermodynamical sense, rather a parameter in the distribution function, which drops as $T \propto a^{-1}$.

1.3 Cosmological observations

It is only in comparatively recent times that cosmology has become a data driven science, in which theoretical hypothesis can be falsified or validated against observational data. It is amazing that only 15 years ago the total energy density of the universe was known with order-of-magnitude accuracy only. Nowadays, most cosmological parameters are constrained within a few percent. The discovery and accurate mapping of CMB fluctuations has constituted a major pillar in this evolution and represents a fundamental cornerstone of modern cosmology, see § 5.3 for an overview.

It is nevertheless of equal importance that many other cosmological probes have been developed in parallel, and this for at least two good reasons. Firstly, all observation suffers in one form or in another from the *degeneracy problem*: only a certain combination of cosmological parameters can be measured accurately. Since degeneracy directions are different for different observations, combining two or more measurements leads to tighter constrains on the parameters we are interested in. The second reason is that cosmologically relevant measurements are intrinsically difficult. One obvious obstacle is that there is only one universe for which the experimental conditions cannot be manipulated at will. Very often the

Quantity		Value	Observations
Baryon density	ω_b	0.024	CMB, BBN, light elements abundance
Cold dark matter density	$\omega_{\rm cdm}$	0.116	CMB+LSS+SN, clusters
Λ density	ω_Λ	0.378	CMB+LSS+SN+weak lensing
Hubble constant	h	0.72	HST, SZ, strong lensing
Optical depth	$\tau_{\rm re}$	0.17	CMB
Spectral index	$n_{\rm s}$	1.00	CMB, LSS, Lyman-α, clusters
Baryons	Ω_b	0.046	
Cold dark matter	$\Omega_{\rm cdm}$	0.224	
Cosmological constant	Ω_Λ	0.73	
Radiation	$\Omega_{\rm rad}$	$7.95 \cdot 10^{-5}$	CMB
Massless ν families	N_ν	3.04	CMB+LSS
Curvature	$\Omega_{\mathcal{K}}$	0.00	CMB+LSS+SN+weak lensing
Initial conditions	purely adiabatic		CMB

Table 1.1: Parameters of today's "ΛCDM cosmological concordance model", which is in good agreement with most of the current observational evidence coming from CMB (Spergel et al., 2003), large scale structures (LSS) (Tegmark et al., 2004b), Big-Bang Nucleosynthesis (BBN) (Fields & Sarkar, 2004), supernovæ type Ia (SN) (Tonry et al., 2003), strong (Kochanek & Schechter, 2004) and weak lensing (Contaldi et al., 2003), Lyman-α absorption systems (Seljak et al., 2003a) and galaxy clusters (Bahcall et al., 2003) observations.

interesting physics is hidden behind foreground emissions, poor statistical sampling, faint signals and non-linearities. It is common to try and extract cosmological information by using objects whose physical properties are poorly understood, and in general systematics are very difficult to assess in cosmology. Hence a cosmological measurement is usually considered as valid only if confirmed by one or more independent pieces of evidence.

The so-called ΛCDM *concordance model* is strongly supported by several independent observational data. It is generally accepted that our universe is very close to flat ($\Omega_{\mathcal{K}} \approx 0$); that it is dominated by "dark energy" ($\Omega_\Lambda \approx 0.7$), perhaps in form of vacuum energy, or quintessence or a tracking scalar field; that around 25% is non-interacting cold dark matter, and that only the remaining 5% is constituted of baryons. If the three neutrino families of the Standard Model of particle physics are not massless (as the large mixing angle solution to the solar neutrino problem seems to suggest), than their mass is bounded from above to be $m_\nu \lesssim \mathcal{O}(1)$eV. Structure formation proceeded by gravitational instability from quantum fluctuations stretched to super-horizon scale by a period of superluminal expansion (inflation). The simplest inflationary model, in which inflation is driven by one single slow-rolling scalar field, successfully predicts the absence of non-Gaussianity, the (predominantly) adiabatic nature of the fluctuations and the almost scale invariant spectral index ($n_{\rm s} \sim 1$) for the perturbations. The age of the universe, around 13 Gyrs, easily accommodates the oldest observed objects. For definiteness, in Table 1.1 we give the parameters of what we believe is a currently widely accepted "concordance model", to which we will refer throughout this work for illustrative and comparative purposes.

Apart from CMB anisotropies, which we will discuss in depth in the rest of this work, we

briefly present some of the pieces of observational evidence which corroborate the (presently) standard ΛCDM scenario.

1.3.1 Big-Bang Nucleosynthesis

Big-Bang Nucleosynthesis is based on the Standard Model of particle physics, and gives predictions for the abundance of light elements D, ^{3}He, ^{4}He and ^{7}Li synthesized in the early Universe, which are in good overall agreement with the observed abundances, see Olive et al. (2000) for a review and Fields & Sarkar (2004) for more recent results.

Below a temperature $T \sim 1$ MeV the neutron-proton conversion rate falls below the expansion rate, and the neutron to proton ratio freezes out at the value $n/p = \exp(-Q/T) \approx 1/6$, where $Q = 1.293$ MeV is the neutron-proton mass difference. The light elements production starts slightly afterwards, at a temperature $T \sim 0.1$ MeV, which is well below the binding energy of deuterium, $B_D = 2.23$ MeV because photo-dissociation prevents the formation of deuterium and other nuclei until then. By this time, β-decay has further reduced the neutron-to-proton ratio to $n/p \approx 1/7$. The surviving neutrons end up almost completely in ^{4}He, while the abundance of the other elements is sensitively dependent on the nuclear reactions rates, which in turn depend on the baryon density, usually expressed with respect to the photon density by defining the parameter η_{10} as

$$\eta_{10} \equiv \frac{n_b}{n_\gamma} \times 10^{10} \approx 274 \cdot \omega_b(\eta_0)\,, \tag{1.47}$$

where η_0 is the conformal time today. A simple counting argument, see Eq. (6.16, page 148), yields that the primordial ^{4}He mass fraction is about 25%, while the number densities of the other elements relative to hydrogen turn out to be of the order D/H $\sim$ ^{3}He/H $\sim 10^{-5}$ and ^{7}Li/H $\sim 10^{-10}$. The predictions are very reliable and accurate, with a residual numerical uncertainty which depends on the experimentally determined reaction rates; interestingly, it turns out that most of this uncertainty is associated with our only approximative knowledge of the neutron lifetime (Cuoco et al., 2003). The other free parameter of BBN is the radiation density in the early Universe, which sets the Hubble expansion rate and therefore determines the freeze-out temperature for the weak reactions and is usually parameterized with the equivalent number of (massless) neutrino families. We comment on the possibility of a non-standard number of neutrino families and discuss BBN-related issues in § 6.1.2.

In summary, agreement between the abundance of the light elements as inferred from astrophysical measurement and the corresponding prediction of BBN is a powerful tool to verify the Standard Model of particle physics. In § 6.2.3 we present in detail the determination of light elements, discuss the slight discrepancies between them and the BBN predictions and give some possible interpretations. However, the overall agreement is satisfactory, and (for a standard number of neutrino families) the light elements abundances can be explained by a baryon density compatible with the one independently inferred from CMB, namely $\eta_{10} \sim 5.5$ or $\omega_b \sim 0.02$.

1.3.2 Matter distribution

Structure formation proceeds from small inhomogeneities in the matter distribution which grow by gravitational instability, eventually giving rise to the large scale structures like

galaxies and clusters observed today. From the determination of the statistical distribution of matter one tries to reconstruct the properties of the primeval fluctuations, and to validate the structure formation model.

In § 3.6.3 we introduce the linear matter power spectrum $P_m(k)$, which represents the Fourier transform of the 2-point correlation function for the matter density contrast. Observations of the distribution of galaxies out to a redshift $z \sim 0.1$ probe the galaxy-galaxy power spectrum, P_{gg}; the Sloan Digital Sky Survey, for example, currently contains approximately 2×10^5 galaxies (Tegmark et al., 2004a), and upon completion will achieve 10^6 galaxies. The problem is then to relate $P_{gg}(k)$, which probes the luminous matter distribution, with the underlying $P_m(k)$ describing (mostly) the dark matter distribution. This is the issue of *bias*, introduced by Kaiser to explain the different amplitudes of the correlation function for galaxies and for clusters (Kaiser, 1984, 1987): the basic idea is that galaxies represent peaks of the matter distribution, and therefore our observations of P_{gg} actually select only the regions of the underlying matter distribution above some threshold. This concept has been extended to various kinds of bias: luminosity-dependent, morphology-dependent, color-dependent bias, scale-dependent bias, anti-bias, and others. The simplest form is to assume a scale-independent bias, which seems to be justified on large (linear) scales, setting

$$P_{gg}(k) = b^2 P_m(k) \quad \text{for } k < k_{\rm NL} \approx 0.3\, h{\rm Mpc}^{-1} \tag{1.48}$$

with the bias parameter b which is just an unknown constant factor (see however e.g. Durrer et al., 2003a for a critical discussion). In practice, this prescription amounts to introducing a free parameter which controls the amplitude of the matter power spectrum. There are methods which allow to determine the bias from the higher-order n-point function of the distribution: for instance Verde et al. (2002) found $b = 1.04 \pm 0.11$ from the data of the 2dF Galaxy Redshift survey (Colless et al., 2001), which plans to measure 2.5×10^5 galaxies.

One can also consider the distribution of galaxy clusters as a function of redshift, which in principle one should be able to predict by using hydro-dynamical simulations. Comparison with the observed distribution would then allow to constrain the cosmological parameters. This simple sounding program is in practice complicated by the need of accurately simulating all the relevant physics, and despite the great amount of computational power nowadays available, recent works in the field still involve many approximations. As a result, cluster data mainly constrain a combination of the matter power spectrum at clusters scales and the value of Ω_m, see e.g. Bahcall et al. (2003).

Another way to probe the mass distribution is offered by the Lyman α forest, the absorption lines in the spectra of distant quasars produced by the neutral hydrogen in regions of overdense intergalactic gas along the line of sight at a redshift $2-4$ (Croft et al., 2002). Since the overdensities probed at these redshifts are still close to the linear regime, one hopes to be able to connect the observations to the matter power spectrum by modelling numerically the relevant physics (Mandelbaum et al., 2003; Seljak et al., 2003a).

Weak gravitational lensing is very promising as a tool to constrain cosmological parameters, and in particular the matter distribution. It uses the distortion in the images of distant galaxies induced by inhomogeneities in the intervening matter distribution (Kaiser & Squires, 1993), and reconstructs with a statistical analysis the so-called "cosmic shear" (Wittman et al., 2000; Bartelmann & Schneider, 2001). The technique is now rapidly becoming mature to help constrain the matter budget (Contaldi et al., 2003).

One of the most important aspects is that all of the above observations can be combined to achieve superior constraining power on the CDM model parameters, while testing the consistency of the theory itself, or the soundness of each data-set. A technique to merge galaxy surveys, cluster distribution, weak lensing and Lyman α data with the CMB to probe a larger portion of the matter power spectrum is presented in Tegmark & Zaldarriaga (2002). There is presently a general agreement that the matter content of the Universe is low, around $\Omega_m \sim 0.3$.

1.3.3 Type Ia supernovæ

Supernovæ (SN) are classified according to their spectrum: the type Ia is characterized by the absence of hydrogen (the "I"), and by strong silicon features (the "a"). The standard picture is a progenitor binary system, with a white dwarf which accretes matter from its companion until it reaches the Chandrasekhar limit, and the gravitational infall triggers a thermonuclear explosion which we observe as a supernova. At the peak of its brightness, a SN can easily exceed the luminosity of its host galaxy, making it a promising candidate to measure distances out to very high ($z \sim 1-2$) redshifts.

Their most important property is the remarkable homogeneity in their spectra, in the shape of their light-curve and in their peak absolute magnitude, which makes them *nearly* "standard candles". In fact, it was discovered that intrinsically brighter SNIa decline more slowly than dim ones (Hamuy et al., 1996). By exploiting an empirical correlation between the shape of the light curve and the intrinsic luminosity, and correcting for extinction effects via measurements at different wavelengths, it is nevertheless possible to produce a "calibrated candle", with a very narrow peak magnitude dispersion (Riess et al., 1996). For a review of the cosmological applications, see e.g. Filippenko (2004).

The measured apparent magnitude m is related to the absolute magnitude M via the *luminosity distance* D_L

$$m = M + 5\log\left[H_0 D_L(z, \Omega_m, \Omega_\Lambda)\right] + K \tag{1.49}$$

where the "K-correction" compensates for the difference in wavelength of the emitted and received photons due to the expansion, and the luminosity distance of an object at redshift z is defined in terms of the intrinsic luminosity L and of the measured flux ℓ as

$$D_L(z) \equiv \left(\frac{L}{4\pi\ell}\right)^{1/2} . \tag{1.50}$$

The luminosity distance is related to the angular diameter distance by $D_L(z) = (1+z)^2 D_A(z)$. Supernovæ essentially measure the angular diameter distance over a redshift range of $z \sim 0.5-2$, much lower than range probed by the CMB. At such low redshift, the radiation content is negligible, and with $\Omega_\mathcal{K} = 1 - \Omega_m - \Omega_\Lambda$ we obtain from (1.32) and (1.34, page 27)

$$\begin{aligned} H_0 D_L(z_1, \Omega_m, \Omega_\Lambda) = & \frac{1+z_1}{\sqrt{|\Omega_\mathcal{K}|}} \times \\ & \chi\left(\frac{1+z_1}{\sqrt{|\Omega_\mathcal{K}|}} \int_0^{z_1} \left[(1+z)^2(1+z\Omega_m) - \Omega_\Lambda z(2+z)\right]^{-1/2} \mathrm{d}z\right) , \end{aligned} \tag{1.51}$$

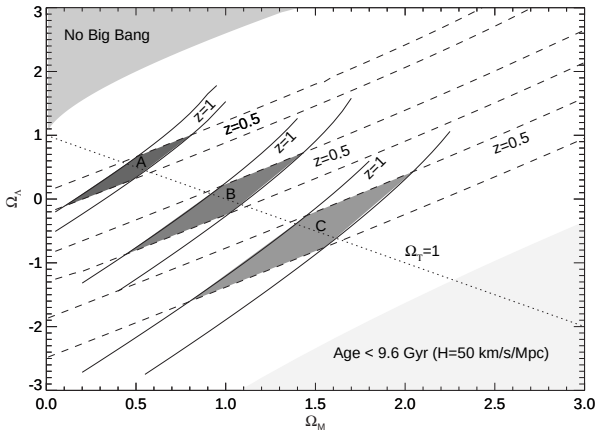

Figure 1.1: Illustration of the determination of $(\Omega_m, \Omega_\Lambda)$ using supernovæ data: the dashed (solid) curves are lines of constant D_L for the given measured apparent magnitude of a standard candle at a redshift $z = 0.5$ ($z = 1.0$). If the apparent magnitude m can be measured with accuracy $\Delta m = 0.05$ combining the two observations gives the dark shaded allowed region for $(\Omega_m, \Omega_\Lambda)$. Figure reprinted from Goobar & Perlmutter (1995).

where the function χ is defined in Eq. (1.3, page 24). Notice that magnitude-redshift relation (1.49) does not depend on the Hubble parameter. Therefore, assuming that we are able to reliably reconstruct the intrinsic luminosity M, from the measurement of one SN Eq. (1.49) yields one degeneracy line for the possible values of $(\Omega_m, \Omega_\Lambda)$. By measuring a second standard candle at $z_2 \neq z_1$ we are able to determine the intersection of the degenerate luminosity distance lines in the $(\Omega_m, \Omega_\Lambda)$ plane, and thus to measure separately the matter and cosmological constant content. When we add the measurements error, both lines widen to two strips, and we obtain a region of confidence for the two parameters, independently on the Hubble parameter, see Fig. 1.1.

In practice, of course, a larger number of measurements is necessary, and it turns out that the approximate combination $\Omega_m - \Omega_\Lambda$ is well constrained, as it is intuitively clear from Fig. 1.1. For instance, Tonry et al. (2003) found

$$\Omega_\Lambda - 1.4\Omega_m = 0.35 \pm 0.14 \quad (\text{at } 1\sigma). \tag{1.52}$$

This degeneracy direction is almost orthogonal to the one in inferred from the angular diameter distance at $z \sim 1100$ measured by the CMB, cf. Fig. 4.1. Combination of supernovæ and CMB data is thus a very effective way to break the angular diameter distance degeneracy and to constraint the matter and vacuum energy contents separately. As we have seen, observations of the matter distribution on large scales independently constrain the matter density parameter: it is a remarkable achievement of modern cosmology that this "cosmic complementarity" seems to be pointing toward the same value, namely $\Omega_m \sim 0.3$ and $\Omega_\Lambda \sim 0.7$, see e.g. Spergel et al. (2003). At the same time, the puzzle of the nature of dark matter and dark energy remains unsolved, and we offer some further remarks regarding the cosmological constant in § 7.3.

Chapter 2

Cosmological perturbation theory

In order to understand the physical origin of CMB anisotropies, we are interested in studying the evolution of perturbations in the photon distribution function, by perturbing at linear order around the "background" solution for the homogeneous and isotropic Friedmann-Robertson-Walker (FRW) universe of § 1.2. That linear perturbation theory is sufficient to describe almost all aspects of CMB physics is a consequence of the smallness of the fluctuations.

In § 2.1 we introduce the relevant perturbation variables, discuss the issues of gauge transformations and gauge invariant formalism, extend the treatment to multiple fluids and define entropy perturbations. We then present the perturbed Einstein (§ 2.2.1) and conservation equations (§ 2.2.2) for an Universe filled with four different particle species: baryons, cold dark matter (CDM), photons and massless neutrinos. The Bardeen equation is presented in § 2.2.3, while § 2.2.4 is devoted to the derivation of the collisionless Boltzmann equation, which describes massless neutrinos and photons after decoupling. The last section § 2.2.5 concerns the Thomson scattering process which couples photons and baryons before recombination, and explains the origin of CMB polarization.

Cosmological perturbation theory in the four-dimensional FRW universe is a well studied subject, see e.g. Kodama & Sasaki (1984); Mukhanov et al. (1992); Ma & Bertschinger (1995); Durrer (1994). More recently, the formalism has been extended to higher-dimensional manifolds, involving extra dimensions (see e.g. Riazuelo et al., 2002), in view of the recent interest in string theory motivated braneworlds scenarios.

2.1 Perturbation variables

In this section, background (unperturbed) quantities are denoted by an overline, so that the perturbed energy density, e.g., is denoted by $\rho = \bar{\rho} + \delta\rho$. The background quantities depend on time only, while the linear perturbations are functions of time and of the 3-space coordinate, i.e. $\delta\rho = \delta\rho(\eta, \mathbf{x})$.

2.1.1 Metric perturbations

We perturb to linear order the FRW metric of Eq. (1.4, page 24) by setting

$$g_{\mu\nu}\mathrm{d}x^\mu \mathrm{d}x^\nu = \bar{g}_{\mu\nu}\mathrm{d}x^\mu \mathrm{d}x^\nu + a^2 h_{\mu\nu}\mathrm{d}x^\mu \mathrm{d}x^\nu \tag{2.1}$$

where the perturbation $h_{\mu\nu}$ is given by

$$h_{\mu\nu}\mathrm{d}x^\mu\mathrm{d}x^\nu = -2A\mathrm{d}\eta^2 + 2B_i\mathrm{d}x^i\mathrm{d}\eta + 2H_{ij}\mathrm{d}x^i\mathrm{d}x^j\,. \tag{2.2}$$

The perturbation variables A, B_i, H_{ij} are arbitrary functions of the 4-coordinate vector $x^\mu = (\eta, \mathbf{x})$.

It is convenient to split them into components which transform irreducibly under the rotation group $SO(3)$. The vector field B_i can thus be written as the sum of a gradient of a scalar and a divergenceless component (vector)

$$B_i = B_{|i} + B_i^{(\mathrm{v})}\,, \quad B_i^{(\mathrm{v})|i} = 0\,. \tag{2.3}$$

We split H_{ij} into an isotropic and an anisotropic part

$$H_{ij} = C\bar{\gamma}_{ij} + E_{ij}\,, \tag{2.4}$$

and E_{ij} is further decomposed in irreducible scalar (spin 0), vector (spin 1) and tensor (spin 2) components as

$$E_{ij} = E_{|ij} + \frac{1}{2}(E_{j|i}^{(\mathrm{v})} + E_{i|j}^{(\mathrm{v})}) + E_{ij}^{(\mathrm{t})}\,, \tag{2.5}$$

where

$$E_{|j}^{(\mathrm{v})j} = E_{|j}^{(\mathrm{t})ij} = 0 \quad \text{(divergenceless)}\,, \tag{2.6}$$

$$E_j^{(\mathrm{t})j} = 0 \quad \text{(traceless)}\,. \tag{2.7}$$

Note that at this stage we are still working in real space and we do not perform an harmonic analysis of the perturbation variables (see Kodama & Sasaki, 1984; Durrer, 1994 instead). At linear order, the different spin components do not mix, and we can treat them separately.

2.1.2 Perturbations of the energy-momentum tensor

The perturbed energy-momentum tensor is obtained by perturbing in Eq. (1.5) the energy density

$$\rho = \bar{\rho} + \delta\rho = \bar{\rho}(1+\delta)\,, \quad \text{with} \quad \delta \equiv \delta\rho/\bar{\rho}\,, \tag{2.8}$$

the pressure

$$P = \bar{P} + \delta P \equiv \bar{P}(1+\pi_L)\,, \quad \text{with} \quad \pi_L \equiv \delta P/\bar{P}\,, \tag{2.9}$$

and the space components of the observer's 4-velocity

$$u^i = \delta u^i \equiv -\frac{v^i}{a} = -\frac{1}{a}(v^{|i} + v^{(\mathrm{v})i})\,, \tag{2.10}$$

$$u^0 = \bar{u}^0 + \delta u^0 = \frac{1}{a}(1-A)\,, \tag{2.11}$$

and the second line follows from the norm of the 4-velocity $u^\mu u_\mu = -1$.

The perturbation of the energy-momentum tensor is then written as

$$\delta T_{\mu\nu} = \left(\bar{\rho}\delta + \bar{P}\pi_L\right)\bar{u}_\mu\bar{u}_\nu + \left(\bar{\rho}+\bar{P}\right)\left(\delta u_\mu\bar{u}_\nu + \delta u_\nu\bar{u}_\mu\right) + \bar{P}\left(\pi_L\bar{g}_{\mu\nu} + a^2h_{\mu\nu} + a^2\Pi_{\mu\nu}\right)\,, \tag{2.12}$$

where we have introduced the *anisotropic stress perturbation* $\Pi_{\mu\nu}$, which is a traceless tensor and orthogonal to the 4-velocity, $u^\mu \Pi_{\mu\nu} = 0$. It describes off-diagonal, space-space perturbations in the stress-energy tensor, and can be split into a scalar Π, a divergenceless vector $\Pi_i^{(\mathrm{v})}$ and a trace-free tensor part $\Pi_{ij}^{(\mathrm{t})}$, according to:

$$\Pi_{ij} = (\nabla_i \nabla_j - \frac{1}{3}\bar{\gamma}_{ij}\nabla_k \nabla^k)\Pi + \frac{1}{2}(\Pi^{(\mathrm{v})}_{i|j} + \Pi^{(\mathrm{v})}_{j|i}) + \Pi^{(\mathrm{t})}_{ij}\,, \tag{2.13}$$

The perturbation components of the stress-energy tensor therefore take the form

$$\delta T^0{}_0 = -\bar{\rho}\delta\,, \tag{2.14a}$$
$$\delta T^0{}_i = (\bar{\rho} + \bar{P})(B_i - v_i)\,, \tag{2.14b}$$
$$\delta T^i{}_0 = (\bar{\rho} + \bar{P})v^i\,, \tag{2.14c}$$
$$\delta T^i{}_j = \bar{P}(\bar{\gamma}^i{}_j \pi_L + \Pi^i{}_j)\,. \tag{2.14d}$$

2.1.3 Gauge transformations

By choosing the background spacetime manifold and metric to be of the FRW form, we implicitly assume that for all quantity of interest Q we are able to define a spatially averaged $\bar{Q}$, which represents the background, homogeneous and isotropic value of Q on $(\bar{\mathcal{M}}, \bar{g})$. Consider now a slightly perturbed manifold, $\mathcal{M}_{\mathrm{pert}}$, endowed with a coordinate system x^μ. The value of Q on $\mathcal{M}_{\mathrm{pert}}$ depends on the choice of the coordinate system, $Q_{\mathrm{pert}} = \bar{Q} + \delta Q(x^\mu)$. Along with x^μ, any other coordinate system which leaves $\bar{g}$ invariant is admissible, i.e. we can arbitrarily transform the coordinates by an infinitesimal amount

$$x^\mu \quad \rightarrow \quad y^\mu = x^\mu + \delta x^\mu \tag{2.15}$$

thereby obtaining for Q in this newly defined coordinates

$$Q_{\mathrm{pert}}(x^\mu) \quad \rightarrow \quad Q_{\mathrm{pert}}(y^\mu) = Q_{\mathrm{pert}}(x^\mu) + \mathcal{L}_{\delta x}(\bar{Q})\,, \tag{2.16}$$

where $\mathcal{L}_X(\bar{Q})$ is the Lie derivative of Q with respect to the vector field X, see e.g. Straumann (2004). Such infinitesimal coordinate transformations are called *gauge transformations*, and the above result is known as Stewart–Walker Lemma. Fixing the coordinate system on $\mathcal{M}_{\mathrm{pert}}$ is called a *gauge choice*. Clearly, physical observables are geometrical quantities, and are therefore independent of the coordinate system in which they are calculated. The form of the equations, however, can be very different according to the gauge choice. It is often convenient to fix the gauge in the way which is best suited for the problem at hand.

The gauge transformation Eq. (2.15) can be written in all generality as

$$\delta x^0 = T\,, \quad \delta x^i = L^{|i} + L^{(\mathrm{v})i}\,. \tag{2.17}$$

By applying the transformation law (2.16) to the perturbed metric (2.1) under a gauge transformation of the type (2.17), we obtain the following transformation properties for the

metric variables:

$$A \rightarrow A + \mathcal{H}T + \dot{T}\,, \tag{2.18a}$$
$$B \rightarrow B - T + \dot{L}\,, \tag{2.18b}$$
$$C \rightarrow C + \mathcal{H}T\,, \tag{2.18c}$$
$$E \rightarrow E + L\,, \tag{2.18d}$$
$$B^{(\mathrm{v})i} \rightarrow B^{(\mathrm{v})i} + \dot{L}^{(\mathrm{v})i}\,, \tag{2.18e}$$
$$E^{(\mathrm{v})i} \rightarrow E^{(\mathrm{v})i} + L^{(\mathrm{v})i}\,, \tag{2.18f}$$
$$E^{(\mathrm{t})ij} \rightarrow E^{(\mathrm{t})ij}\,. \tag{2.18g}$$

The same procedure applied on the background stress-energy tensor $\bar{T}_{\mu\nu}$ and 4-velocity $\bar{u}^\mu$ gives for the matter perturbation variables:

$$\delta \rightarrow \delta - 3T\mathcal{H}(1+w)\,, \tag{2.19a}$$
$$\pi_L \rightarrow \pi_L - \frac{3c_s^2}{w}(1+w)\mathcal{H}T\,, \tag{2.19b}$$
$$\Pi \rightarrow \Pi\,, \tag{2.19c}$$
$$v \rightarrow v + \dot{L}\,, \tag{2.19d}$$
$$v_i^{(\mathrm{v})} \rightarrow v_i^{(\mathrm{v})} + \dot{L}_i^{(\mathrm{v})}\,, \tag{2.19e}$$
$$\Pi_i^{(\mathrm{v})} \rightarrow \Pi_i^{(\mathrm{v})}\,, \tag{2.19f}$$
$$\Pi_{ij}^{(\mathrm{t})} \rightarrow \Pi_{ij}^{(\mathrm{t})}\,. \tag{2.19g}$$

In order to completely fix the gauge, we need to specify in Eq. (2.2) the functional form of two scalar functions, corresponding to a specific choice for (T, L), and one vector, corresponding to a choice for $L^{(\mathrm{v})i}$. In the following, we briefly summarize some popular gauge choices.

Longitudinal gauge

Longitudinal gauge (also sometimes called "Newtonian gauge") is defined by requiring $B = E = B^{(\mathrm{v})i} = 0$, so that the perturbed metric element takes the form

$$\mathrm{d}s^2 = a^2\left[-(1+2\Psi)\mathrm{d}\eta^2 + (1-2\Phi)\bar{\gamma}_{ij}\mathrm{d}x^i\mathrm{d}x^j\right]\,, \tag{2.20}$$

and we have defined the *Bardeen potentials* $\Psi = A$ and $\Phi = -C$ (Bardeen, 1980), which represent the gravitational time dilation and the perturbation to the 3-space curvature, respectively. From any other gauge, the transformation $T = B - \dot{E}$, $L = -E$ and $\dot{L}^{(\mathrm{v})i} = -B^{(\mathrm{v})i}$ leads to the longitudinal gauge.

Flat slicing gauge

This gauge owns its name to the choice $E = C = E^{(\mathrm{v})i} = 0$, which makes the spatial hypersurfaces unperturbed, and the metric element is

$$\mathrm{d}s^2 = a^2\left[-(1+2A)\mathrm{d}\eta^2 + 2B_i\mathrm{d}x^i\mathrm{d}\eta + \bar{\gamma}_{ij}\mathrm{d}x^i\mathrm{d}x^j\right]\,. \tag{2.21}$$

The coordinate transformation which leads to flat slicing gauge is $T = -C/\mathcal{H}$, $L = -E$ and $L^{(\mathrm{v})i} = -E^{(\mathrm{v})i}$.

Synchronous gauge

In synchronous gauge, constant time hypersurfaces are orthogonal to the 3-space (hence the name), i.e. (η, x^i) are Gaussian coordinates. This can be obtained by imposing $A = B = B^{(\mathrm{v})i} = 0$. Thus the metric presents perturbations in the space-space part only, and it is often written as

$$\mathrm{d}s^2 = a^2\left[-\mathrm{d}\eta^2 + (\bar{\gamma}_{ij} + h_{ij})\mathrm{d}x^i\mathrm{d}x^j\right], \tag{2.22a}$$

$$h_{ij} \equiv h_{|ij}(\eta, \mathbf{x}) + (\nabla_i\nabla_j - \frac{1}{3}\bar{\gamma}_{ij}\nabla_k\nabla^k)6\eta(\eta, \mathbf{x})\,. \tag{2.22b}$$

The above choice does not fix completely the gauge: in fact, the gauge transformation which leads to synchronous gauge is

$$T = -\frac{1}{a}\int aA\mathrm{d}\eta + \frac{\alpha}{a} \tag{2.23a}$$

$$L = \int (T - B)\mathrm{d}\eta + \beta \tag{2.23b}$$

$$L^{(\mathrm{v})i} = -\int B^{(\mathrm{v})i}\mathrm{d}\eta + \beta^{(\mathrm{v})i}\,, \tag{2.23c}$$

which presents a residual gauge freedom in the four arbitrary integration constants α and $\beta^i = \beta^{|i} + \beta^{(\mathrm{v})i}$ (where $\beta^{(\mathrm{v})i}$ must be divergenceless). The four constants correspond to different choices of the constant time hypersurface and of the spatial coordinates on it. This leads to the presence of fictitious "gauge modes" in the perturbation equations, which must be removed because they are just an artifact of the choice of the coordinate. Despite this difficulty, synchronous gauge is quite popular in the literature.

Comoving gauge

In the comoving gauge the total bulk velocity vanishes, $\delta T^0_{\ i} = 0$, which translates into the condition $B_i = v_i$. In order to completely fix the gauge one further requires $E = 0$ and $E^{(\mathrm{v})i} = 0$. This is achieved with the transformation $T = B - v - \dot{E}$, $L = -E$ and $L^{(\mathrm{v})i} = -E^{(\mathrm{v})i}$. This gauge is the one which resembles most the gauge invariant formalism (defined below), since for the variables in comoving gauge we have

$$\begin{aligned} C &= -\zeta \quad && \text{see Eq. (2.26)} \\ \delta &= D \quad && \text{see Eq. (2.30)} \\ \delta_\alpha &= \Delta_\alpha \quad && \text{see Eq. (2.37)} \\ v &= V \quad && \text{see Eq. (2.31)}\,. \end{aligned} \tag{2.24}$$

2.1.4 Gauge invariance

General covariance guarantees that all equations in general relativity can be written in a form which is independent of the gauge choice (Bardeen, 1980; Kodama & Sasaki, 1984; Durrer, 1994). From (2.16) it follows that for all tensor fields with vanishing or constant background contribution, so that $\mathcal{L}_X(\bar{Q}) = 0 \quad \forall X$, we can construct *gauge invariant perturbation equations*. Such perturbation variables are invariant under a gauge transformation of the type

Eq. (2.15). Since we can cast all general relativistic equations in the form $Q = 0$, it is always possible to construct gauge invariant perturbation equations (Stewart & Walker, 1974).

This approach has the advantage of leading to equations which are independent of the coordinate choice, and which are often easier to interpret physically. Furthermore, gauge independent equations are free from spurious gauge modes. In order to write down the relevant gauge invariant perturbation equations, we make use of the transformation properties of the metric and matter variables under a change of gauge, Eqs. (2.18) and (2.19).

Metric variables

From Eq. (2.18) we can construct the following 4 gauge invariant metric variables:

$$\Phi \equiv -C - \mathcal{H}(B - \dot{E})\,, \tag{2.25a}$$

$$\Psi \equiv A + \mathcal{H}(B - \dot{E}) + (\dot{B} - \ddot{E})\,, \tag{2.25b}$$

$$\Sigma_i^{(\mathrm{v})} \equiv \dot{E}_i^{(\mathrm{v})} - B_i^{(\mathrm{v})}\,, \tag{2.25c}$$

$$H_{ij}^{(\mathrm{t})} \equiv E_{ij}^{(\mathrm{t})}\,. \tag{2.25d}$$

The two scalar variables Φ and Φ are called Bardeen potentials (Bardeen, 1980). Another very useful variable is the *gauge invariant curvature perturbation* ζ, which is defined as

$$\zeta \equiv -C + \mathcal{H}(v - B)\,, \tag{2.26}$$

where v is defined in Eq. (2.10). From the constraint equation (2.50), it follows that for a flat universe, $\mathcal{K} = 0$, the gauge invariant curvature perturbation is related to the Bardeen potentials by

$$\zeta = \Phi + \frac{\mathcal{H}}{\mathcal{H}^2 - \dot{\mathcal{H}}}(\mathcal{H}\Psi + \dot{\Phi})\,. \tag{2.27}$$

There is only one gauge invariant vector perturbation constructed out of metric variables, Eq. (2.25c). Tensor variables are automatically gauge invariant, since there is no spin-2 coordinate transformation.

Matter variables

Because of the Stewart–Walker Lemma (2.16), the variables Π, $\Pi_i^{(\mathrm{v})}$ and $\Pi_i^{(\mathrm{t})}$ are already gauge invariant, since the background anisotropic stress vanishes.

From scalar matter variables alone we can construct the gauge invariant variable

$$\Gamma \equiv \pi_L - \frac{c_s^2}{w}\delta\,, \tag{2.28}$$

which measures the intrinsic non-adiabaticity of the matter content. More precisely, as we shall see below, Γ is related to the entropy production rate. If the pressure is a function of the local energy density only, $P = P(\rho)$, then we can write

$$\frac{\delta P}{\delta \rho} = \frac{\dot{P}}{\dot{\rho}} \tag{2.29}$$

and since by definition $\delta\rho = \delta \cdot \rho$, $\delta P = \pi_L \cdot P$, it follows that $\Gamma = 0$. In the case of a perfect fluid, $P = w\rho$ and Γ vanishes. Non-zero contributions to Γ arise from the relative entropy of a mixture of several fluid components, which is discussed in § 2.1.5.

The choice of a gauge invariant density contrast is not unique, and requires the use of metric variables. Meaningful combinations are

$$D_s \equiv \delta - 3(1+w)\mathcal{H}(B - \dot{E}) \quad \text{(longitudinal)}, \tag{2.30a}$$

$$D_g \equiv \delta + 3(1+w)C \quad \text{(flat slicing)}, \tag{2.30b}$$

$$D \equiv \delta - 3(1+w)\mathcal{H}(B - v) \quad \text{(comoving)}. \tag{2.30c}$$

On super-horizon scales, D_s corresponds to the density contrast in the longitudinal gauge; D_g is the density contrast on homogeneous 3-space hypersurfaces (flat slicing); D reduces to the density contrast in the comoving gauge. The distinction is only important on super-horizon scales, since on small (sub-horizon) scales, all the above variables reduce to the same (Durrer, 2001).

The remaining velocity perturbation can be written in gauge invariant form as

$$V \equiv v - \dot{E}\,, \tag{2.31a}$$

$$V_i^{(\mathrm{v})} \equiv v_i^{(\mathrm{v})} - \dot{E}_i^{(\mathrm{v})}\,. \tag{2.31b}$$

Useful relations between those gauge invariant variables are

$$D_g = D_s - 3(1+w)\Phi\,, \tag{2.32a}$$

$$D = D_s + 3(1+w)\mathcal{H}V\,, \tag{2.32b}$$

$$D = D_g + 3(1+w)\zeta\,, \tag{2.32c}$$

$$\zeta = \Phi + \mathcal{H}V\,. \tag{2.32d}$$

2.1.5 Multiple fluids

The above definitions assume that the universe is filled with, or dominated by, only one fluid component. In a more realistic modelling, we must account for the presence of several matter components. We will usually consider four of them, namely photons (subscript γ), massless neutrinos (subscript ν), non-interacting cold dark matter (CDM, subscript c) and baryons (subscript b). The subscripts r (radiation) and m (matter) will refer generically to a relativistic ($w_r = 1/3$) and a non-relativistic, dust-like ($w_m = 0$) fluid, respectively. Variables without subscript designate the total perturbation.

If multiple matter components are present, the total perturbation variables are the weighted sum of the variables for each component:

$$\delta = \sum_\alpha \frac{\bar{\rho}_\alpha}{\bar{\rho}}\delta_\alpha\,, \tag{2.33a}$$

$$v^j = \sum_\alpha \frac{\bar{\rho}_\alpha + \bar{P}_\alpha}{\bar{\rho} + \bar{P}} v_\alpha^j\,, \tag{2.33b}$$

$$\Pi^{ij} = \sum_\alpha \frac{\bar{P}_\alpha}{\bar{P}}\Pi_\alpha^{ij}\,. \tag{2.33c}$$

The equation of state and the adiabatic sound speed are defined for each component

$$w_\alpha \equiv \frac{\bar{P}_\alpha}{\bar{\rho}_\alpha} \quad \text{and} \quad c_\alpha^2 \equiv \frac{\dot{\bar{P}}_\alpha}{\dot{\bar{\rho}}_\alpha}\,, \tag{2.34}$$

and for the mixture we have

$$w \equiv \frac{\bar{P}}{\bar{\rho}} \quad \text{and} \quad c_s^2 \equiv \frac{\dot{\bar{P}}}{\dot{\bar{\rho}}} \,. \tag{2.35}$$

The transformation properties of the variables for each components are the same as for the total variables, Eqs. (2.19). Hence for each matter component we can define gauge invariant variables as in Eqs. (2.28, 2.30, 2.31), yielding for the scalar part:

$$\Gamma_\alpha \equiv \pi_{\alpha,L} - \frac{c_\alpha^2}{w_\alpha}\delta_\alpha \,, \tag{2.36a}$$

$$V_\alpha \equiv v_\alpha - \dot{E} \,, \tag{2.36b}$$

$$D_{\alpha,s} \equiv \delta_\alpha - 3(1+w_\alpha)\mathcal{H}(B-\dot{E}) \,, \tag{2.36c}$$

$$D_{\alpha,g} \equiv \delta_\alpha + 3(1+w_\alpha)C \,, \tag{2.36d}$$

$$D_\alpha \equiv \delta_\alpha - 3(1+w_\alpha)\mathcal{H}(B-v_\alpha) \,. \tag{2.36e}$$

In the presence of multiple matter components, it is often useful to work with the gauge invariant density contrast

$$\Delta_\alpha \equiv \delta_\alpha - 3(1+w_\alpha)\mathcal{H}(B-v) \,, \tag{2.37}$$

which corresponds to the density contrast in the gauge where the total matter is at rest, i.e. the comoving gauge introduced on page 39. Notice that on the right hand side it appears the total velocity v, rather then the velocity of the α component as in (2.36e). This new variable is related to the density contrast in the flat slicing gauge by

$$\Delta_\alpha = D_{g,\alpha} + 3(1+w_\alpha)\left(\Phi + \mathcal{H}V\right) \,. \tag{2.38}$$

2.1.6 Entropy perturbations

When more than one component is present, entropy perturbations can arise even for a mixture of perfect fluids. The total non-adiabaticity of the mixture is given by (2.28), where the quantities appearing on the right hand side have to be interpreted as total variables. Using the definitions (2.33), we obtain

$$\bar{P}\Gamma = \bar{P}\Gamma_{\rm int} + \sum_\alpha \delta_\alpha \bar{\rho}_\alpha (c_\alpha^2 - c_s^2) \,, = \bar{P}\left(\Gamma_{\rm int} + \Gamma_{\rm rel}\right) \,. \tag{2.39}$$

We have introduced the total intrinsic entropy perturbation

$$\Gamma_{\rm int} = \sum_\alpha \frac{\bar{P}_\alpha}{\bar{P}}\Gamma_\alpha \tag{2.40}$$

and the relative entropy perturbation $\Gamma_{\rm rel}$, which using the background energy conservation, Eq. (1.16, page 25), can be recast as

$$\bar{P}\Gamma_{\rm rel} = \frac{1}{2}\sum_{\alpha,\beta} \frac{(1+w_\alpha)(1+w_\beta)\bar{\rho}_\alpha\bar{\rho}_\beta}{(1+w)\bar{\rho}}(c_\alpha^2 - c_\beta^2)\left(\frac{\delta_\alpha}{1+w_\alpha} - \frac{\delta_\beta}{1+w_\beta}\right) \,. \tag{2.41}$$

Here we have assumed that the components are decoupled from each other, i.e. that $\bar{Q}_\alpha^\nu = 0$ in (1.16, page 25), see (Malik et al., 2003) for a generalization to the case of interacting fluids.

The quantity $\Gamma_{\rm rel}$ represents *relative entropy perturbations* which are produced by the different dynamical behavior of the matter components with different sound speed. The *entropy perturbation* between the components α and β is defined as

$$S_{\alpha\beta} \equiv \frac{\delta_\alpha}{1+w_\alpha} - \frac{\delta_\beta}{1+w_\beta}\,. \tag{2.42}$$

It is easy to see that the entropy perturbations are gauge invariant quantities by substituting the gauge dependent density contrasts on the right hand side with the gauge invariant density variables defined in (2.37), obtaining

$$S_{\alpha\beta} = \frac{\Delta_\alpha}{1+w_\alpha} - \frac{\Delta_\beta}{1+w_\beta}\,. \tag{2.43}$$

In order to clarify the physical meaning of $S_{\alpha\beta}$, consider a mixture of radiation and dust-like matter. We are interested in fluctuations of the number density (per physical volume) ratio of the two species:

$$\delta\left(\frac{n_r}{n_m}\right)/(n_r n_m) = \frac{\delta n_r}{n_r} - \frac{\delta n_m}{n_m}\,. \tag{2.44}$$

Recall that (see e.g. Kolb & Turner, 1990) $n_r \propto s \propto T^3$, with s the radiation entropy per volume, hence

$$\frac{\delta n_r}{n_r} = \frac{\delta s}{s} = 3\frac{\delta T}{T} = \frac{3}{4}\frac{\delta\rho_r}{\bar\rho_r} \tag{2.45}$$

For matter we have

$$\frac{\delta n_m}{n_m} = \frac{\delta\rho_m}{\bar\rho_m}\,, \tag{2.46}$$

and therefore

$$\frac{\delta n_r}{n_r} - \frac{\delta n_m}{n_m} = \frac{\delta_r}{(1+w_r)} - \frac{\delta_m}{(1+w_m)} = S_{rm}\,. \tag{2.47}$$

Thus a non vanishing relative entropy perturbation means that there are spatial inhomogeneities in the relative number density of the the two fluids, which can be understood as a spatial variation in the equation of state. The above results are generalized in § 4.3.

2.2 Perturbation equations

In this section, we write down the first order perturbation equations using the gauge invariant formalism and variables defined above. For completeness, we also give the vector and tensor equations, but in the rest of this work we will concentrate exclusively on the scalar sector.

2.2.1 Einstein equations

The perturbed Einstein equations

$$\delta G_{\mu\nu} = 8\pi G \delta T_{\mu\nu} \tag{2.48}$$

are split in their scalar, vector and tensor parts.

Scalar equations

There are 4 scalar equations for the 4 gauge invariant quantities Φ, Ψ, V and D:

$$(\triangle + 3\mathcal{K})\Phi = 4\pi G a^2 \bar{\rho} D \qquad \text{(Poisson)}, \tag{2.49}$$

$$\mathcal{H}\Psi + \dot{\Phi} = 4\pi G a^2 \bar{\rho}(1+w)V \qquad \text{(constraint)}, \tag{2.50}$$

$$\Phi - \Psi = 8\pi G a^2 \bar{\rho} w \Pi \qquad \text{(anisotropic stress)}, \tag{2.51}$$

$$\mathcal{H}\dot{U} + (\mathcal{H}^2 + 2\dot{\mathcal{H}})U = 4\pi G a^2 \bar{\rho} \left(c_s^2 D_g + w\Gamma + \frac{2}{3} w \triangle \Pi \right), \tag{2.52}$$

where

$$U \equiv \Psi + \frac{\mathcal{H}^2 - \dot{\mathcal{H}}}{\mathcal{H}^2}\Phi + \frac{\dot{\Phi}}{\mathcal{H}}\,. \tag{2.53}$$

Recall that D_g is related to D, V and Φ via Eqs. (2.32, page 41), and we have assumed an equation of state of the form (1.7, page 24). Eq. (2.49) is the general relativistic analogue of the Poisson equation. In order to close this system, we need to specify the matter content by giving w, c_s^2, Γ and Π. For a single perfect fluid, $\Gamma = \Pi = 0$, hence from the anisotropic stress equation (2.51) it follows that $\Psi = \Phi$.

We shall see below that an evolution equation for Π follows e.g. from the kinetic description provided by the Boltzmann equation, see Eq. (2.127, page 53). For multiple fluids, we will also rewrite Γ in terms of the relative entropy perturbations, as in Eq. (2.41).

Vector equations

The vector part yields a constraint and an evolution equation for $V_i^{(\mathrm{v})}$ and $\Sigma_i^{(\mathrm{v})}$:

$$\left(2\mathcal{K} + \triangle + 4(\dot{\mathcal{H}} - \mathcal{H}^2)\right)\Sigma_i^{(\mathrm{v})} = 16\pi G \bar{\rho} a^2 (1+w) V_i^{(\mathrm{v})}\,, \tag{2.54}$$

$$\dot{\Sigma}_i^{(\mathrm{v})} + 2\mathcal{H}\Sigma_i^{(\mathrm{v})} = 8\pi G \bar{\rho} a^2 w \Pi_i^{(\mathrm{v})}\,. \tag{2.55}$$

For a perfect fluid, $\Pi_i^{(\mathrm{v})} = 0$, the above equations give in a flat universe on large scales (such that gradients can be neglected)

$$\Sigma_i^{(\mathrm{v})} = -V_i^{(\mathrm{v})} \propto \frac{1}{a^2}\,. \tag{2.56}$$

Therefore in the absence of active seeds, vector perturbations are always decaying on large scales.

Tensor equation

The tensor part yields an equation describing the gravitational waves. It is the equation of a forced harmonic oscillator, with a damping term due to the expansion of the universe:

$$\ddot{E}_{ij}^{(\mathrm{t})} + 2\mathcal{H}\dot{E}_{ij}^{(\mathrm{t})} + (2\mathcal{K} - \triangle)E_{ij}^{(\mathrm{t})} = 8\pi G \bar{\rho} a^2 \Pi_{ij}^{(\mathrm{t})}\,. \tag{2.57}$$

On super-horizon scales and for zero curvature, the term $\propto E_{ij}^{(\mathrm{t})}$ is negligible. The homogeneous equation in the radiation era, when $\mathcal{H} = \eta^{-1}$, has a decaying solution $E_{ij}^{(\mathrm{t})} \propto \eta^{-1}$ and a constant solution, $E_{ij}^{(\mathrm{t})} = \text{const}$. As a mode enters the horizon, the oscillatory behavior takes over, and the wave propagates with a frequency $k^2 + 2\mathcal{K}$ and is damped as a^{-1}. In the

absence of anisotropic stress and in a flat universe, $\mathcal{K} = 0$, the general solution of (2.57) for $\Pi = 0$, writing $E_{ij}^{(\mathrm{t})} = h(\mathbf{x}, \eta)\varepsilon_{ij}(\mathbf{x})$ and going to Fourier space in a flat universe, is given by

$$h = (k\eta)^{1-q} \left[A j_{q-1}(k\eta) + B n_{q-1}(k\eta) \right] , \tag{2.58}$$

where $j_\nu(x)$ and $n_\nu(x)$ are the Bessel and von Neumann functions of order ν, respectively (see Eqs. 3.10, page 60) and $a \propto \eta^q$.

2.2.2 Conservation equations

The conservation equations, which follow from the contracted Bianchi identity, offer evolution equations which are sometimes of a simpler form and are handy to manipulate. From the perturbed energy conservation equation

$$\delta(\nabla_\mu \bar{T}^{\mu\nu}) = 0 \tag{2.59}$$

we obtain the following equations for a mixture of non-interacting fluids.

Scalar equations

There are two scalar conservation equations, one for the density contrast and the second for the velocity perturbation. In terms of $D_{g,\alpha}$ the conservation equations read:

$$\dot{D}_{g,\alpha} + 3\mathcal{H}(c_\alpha^2 - w_\alpha) D_{g,\alpha} = -3\mathcal{H}\Gamma_\alpha w_\alpha + (1 + w_\alpha)\triangle V_\alpha \,, \tag{2.60}$$

$$\dot{V}_\alpha + (1 - 3c_\alpha^2)\mathcal{H} V_\alpha = \Psi + 3c_\alpha^2 \Phi + \frac{w_\alpha}{1 + w_\alpha} \left(\Gamma_\alpha + \frac{c_\alpha^2}{w_\alpha} D_{g,\alpha} + \frac{2}{3}(\triangle + 3\mathcal{K})\Pi_\alpha \right) . \tag{2.61}$$

Is is sometimes convenient to express the above in terms of the density contrast D_α:

$$\dot{D}_\alpha - 3w_\alpha \mathcal{H} D_\alpha = (\triangle + 3\mathcal{K}) \left[(1 + w_\alpha) V_\alpha + 2\mathcal{H} w_\alpha \Pi_\alpha \right] + 3\frac{1 + w_\alpha}{1 + w}(\mathcal{H}^2 + \mathcal{K})(V - V_\alpha) \,, \tag{2.62}$$

$$\dot{V}_\alpha + \mathcal{H} V_\alpha = \Psi + \frac{c_\alpha^2}{1 + w_\alpha} D_\alpha + \frac{w_\alpha}{1 + w_\alpha} \left(\Gamma_\alpha + \frac{2}{3}(\triangle + 3\mathcal{K})\Pi_\alpha \right) . \tag{2.63}$$

Vector equation

We obtain one evolution equation for the *vorticity* $\Omega_{i\alpha}^{(\mathrm{v})} \equiv \Sigma_{i\alpha}^{(\mathrm{v})} + V_{i\alpha}^{(\mathrm{v})}$:

$$\dot{\Omega}_{i,\alpha}^{(\mathrm{v})} + \mathcal{H}\Omega_{i,\alpha}^{(\mathrm{v})}(1 - 3c_\alpha^2) = \frac{1}{2}\frac{w_\alpha}{1 + w_\alpha}\triangle \Pi_{i,\alpha}^{(\mathrm{v})} \,. \tag{2.64}$$

If the anisotropic stress source term is absent, we can rewrite the above equation as

$$\frac{\mathrm{d}}{\mathrm{d}\eta}(\Omega_{i,\alpha}^{(\mathrm{v})} a^{1-3c_\alpha^2}) = 0 \,, \tag{2.65}$$

hence

$$\Omega_{i,\alpha}^{(\mathrm{v})} \propto a^{3c_\alpha^2 - 1} \,. \tag{2.66}$$

2.2.3 The Bardeen equation

It is often convenient to have an evolution equation for the Bardeen potential in terms of the total matter content. By combining the conservation equation Eq. (2.60) with the Einstein equations (2.49–2.51) we obtain a second order equation, called the Bardeen equation, for Φ:

$$\ddot{\Phi} + 3\mathcal{H}(1+c_s^2)\dot{\Phi} + \left[3(c_s^2-w)\mathcal{H}^2 - (1+3c_s^2)\mathcal{K} - c_s^2\triangle\right]\Phi = g_\Phi \,, \tag{2.67}$$

where the source term g_Φ is generated by the matter anisotropic stress and entropy perturbation:

$$g_\Phi = 8\pi G a^2 P\left[\mathcal{H}\dot{\Pi} + [2\dot{\mathcal{H}} + 3\mathcal{H}^2(1-c_s^2/w)]\Pi + \tfrac{1}{2}\triangle\Pi + \tfrac{1}{2}\Gamma\right] . \tag{2.68}$$

The above equation can be recast in an evolution equation for the gauge invariant curvature perturbation, Eq. (2.27). For hydrodynamical matter, i.e. setting $\Pi = 0$ and for a flat universe ($\mathcal{K} = 0$) we find

$$\dot{\zeta} = \frac{\mathcal{H}}{\mathcal{H}^2 - \dot{\mathcal{H}}}\left[c_s^2\triangle\Phi + \tfrac{3}{2}\mathcal{H}^2 w\Gamma\right] . \tag{2.69}$$

This expression will be used when discussing the evolution of curvature and entropy perturbations.

2.2.4 Collisionless Boltzmann equation

We briefly recall in the following the basics of relativistic kinetic theory, for more details see e.g. de Groot et al. (1980). Consider the phase space given by the the tangent bundle

$$\mathcal{T} \equiv \{(x^\mu, p^\mu) | x^\mu \in \mathcal{M}, p^\mu \in \mathcal{T}_x\} \tag{2.70}$$

where $\mathcal{M}$ is the spacetime manifolds and $\mathcal{T}_x$ its tangent space at the point x^μ. For a particle of mass m, its distribution function $f(x^\mu, p^\mu)$ is defined on the mass-shell

$$\mathcal{P}_m(x^\mu) \equiv \{p^\mu \in \mathcal{T}_x | p_\mu p^\mu = -m^2\} \tag{2.71}$$

The Liouville operator $\mathcal{L}$ is defined on $\mathcal{T}$, and it gives the evolution of $f(x^\mu, p^\mu)$ along the particle world lines, according to the *Boltzmann equation*

$$\mathcal{L}[f] = C[f] \,, \tag{2.72}$$

which states that the rate of change of f is due to the *collision term* $C[f]$. For the purpose of studying relativistic particles such as photons and massless neutrinos, we will treat the case $m = 0$ only. The hereby derived equations will then be applied to the description of neutrinos and of photons after recombination. Further details and the general case for massive particles can be found in e.g. Durrer (1994); Uzan (1998).

We now proceed with perturbing the left hand side of Eq. (2.72). Its background solution was presented in § 1.2.2, and was shown to be of the form $\bar{f} = \bar{f}(ap)$, see Eq. (1.45), where $E^2 = p^2 \equiv p_\mu p_\nu g^{\mu\nu}$. By splitting the distribution function into a background and a perturbed part,

$$f(\eta, x^i, p, n^i) = \bar{f}(\eta, p) + F(\eta, x^i, p, n^i) \tag{2.73}$$

we move to a phase space which differs to linear order from the one of $\bar{f}$. Therefore the choice of F and its transformation properties depend on the isomorphism relating the "background"

and the "perturbed" phase space. By an opportune choice of the isomorphism, it can be shown (Durrer, 1994) that under a gauge transformation F transforms as

$$F \quad \rightarrow \quad F + p\frac{\partial \bar{f}}{\partial p}\left[\mathcal{H}T + n^i T_i\right] . \tag{2.74}$$

It follows that the following variable

$$\mathcal{F} \equiv F - p\frac{\partial \bar{f}}{\partial p}\left[C + n^i(\dot{E}_i - B_i)\right] , \tag{2.75}$$

is gauge invariant. In terms of $\mathcal{F}$, the collisionless Boltzmann equation reads

$$\frac{\partial \mathcal{F}}{\partial \eta} + \frac{\partial \mathcal{F}}{\partial x^i}n^i - p\mathcal{H}\frac{\partial \mathcal{F}}{\partial p} - {}^{(3)}\Gamma^i_{jk}n^j n^k \frac{\partial \mathcal{F}}{\partial n^i} = p\frac{\partial \bar{f}}{\partial p}\left[n^i \partial_i(\Psi + \Phi)\right] , \tag{2.76}$$

and ${}^{(3)}\Gamma^i_{jk}$ are the Christoffel symbols of the background 3-space. The above equation is in manifestly gauge invariant form, and we notice that spatial variations in the Bardeen potential act as source for perturbations in the distribution function.

By integrating this equation over the particle energies, we obtain a differential equation for the *brightness perturbation* $\mathcal{I}$, defined as

$$I = \bar{I}(\eta) + \mathcal{I}(\eta, x^i, n^i) \equiv 4\pi \int_0^\infty \bar{f} p^3 \mathrm{d}p + 4\pi \int_0^\infty \mathcal{F} p^3 \mathrm{d}p \, . \tag{2.77}$$

The brightness represents the energy per unit solid angle as measured by an observer at position x^i. The photon energy is just the monopole of the brightness, i.e.

$$\rho_\gamma = \int \frac{\mathrm{d}\Omega}{4\pi} I \, , \tag{2.78}$$

and therefore $\bar{\rho}_\gamma = \bar{I}$. From Eq. (2.76) we obtain

$$\dot{\mathcal{I}} + \left(n^i\partial_i + 4\mathcal{H} - {}^{(3)}\Gamma^i_{jk}n^j n^k \frac{\partial}{\partial n^i}\right)\mathcal{I} = -4\bar{I}\left[n^i\partial_i(\Psi + \Phi)\right] . \tag{2.79}$$

The above can be rewritten in terms of the *temperature contrast*

$$\Theta(\eta, x^i, n^i) \equiv \frac{\delta T}{T} = \frac{1}{4}\frac{\mathcal{I}}{\bar{I}} \tag{2.80}$$

and using the background energy conservation equation we obtain

$$\dot{\Theta} + \left(n^i\partial_i - {}^{(3)}\Gamma^i_{jk}n^j n^k \frac{\partial}{\partial n^i}\right)\Theta = -n^i\partial_i(\Psi + \Phi) \, . \tag{2.81}$$

This is the Boltzmann equation for relativistic, collisionless particles, which relates gravitational perturbations to temperature fluctuations of their distribution function.

The Boltzmann hierarchy

We now go to Fourier space, and we restrict ourselves to the spatially flat case, $\mathcal{K} = 0$, so that the eigenfunctions of the Laplacian are just plane waves and ${}^{(3)}\Gamma^i_{jk} = 0$ (an harmonic decomposition for non-flat spaces can be found e.g. in Vilenkin & Smorodinskii, 1964; Kodama & Sasaki, 1984), so that for any scalar f

$$f(\eta, \mathbf{x}) = \frac{1}{(2\pi)^{3/2}} \int \mathrm{d}^3\mathbf{k} f(\eta, \mathbf{k}) e^{\imath \mathbf{k}\mathbf{x}} \, , \tag{2.82}$$

and in general we denote the real space f and its harmonic transform with the same symbol. Defining $\mu \equiv n^j k_j/k$ and $k \equiv \sqrt{k_i k^i}$ we obtain from Eq. (2.81)

$$\dot{\Theta} + \imath\mu k\Theta = -\imath\mu k(\Psi + \Phi)\,. \tag{2.83}$$

Assuming that Θ does not depend explicitly on k_i, then the dependence on the photons momentum direction comes in only via μ. In that case $\Theta = \Theta(\eta, k, \mu)$, and we will suppress the explicit time dependence. We now perform an expansion in Legendre polynomials[1]

$$\Theta(\mu, k) = \sum_\ell (2\ell + 1) P_\ell \Theta_\ell\,, \tag{2.84}$$

$$\Theta_\ell(k) \equiv \frac{1}{2}\int_{-1}^{1} \mathrm{d}\mu \Theta(\mu, k) P_\ell(\mu)\,, \tag{2.85}$$

where $P_\ell(x)$ is the Legendre polynomial of order ℓ, which satisfy

$$P_0(x) = 1\,, \tag{2.86}$$

$$P_1(x) = x\,, \tag{2.87}$$

$$P_2(x) = \frac{1}{2}(3x^2 - 1)\,, \tag{2.88}$$

$$(\ell + 1)P_{\ell+1}(x) = (2\ell + 1)xP_\ell(x) - \ell P_{\ell-1}(x)\,. \tag{2.89}$$

From Eq. (2.83) follows an infinite hierarchy of equations for the moments of the Boltzmann equation:

$$\dot{\Theta}_0 + \imath k\Theta_1 = 0\,, \tag{2.90}$$

$$\dot{\Theta}_1 + \frac{1}{3}\imath k\Theta_0 + \frac{2}{3}\imath k\Theta_2 = -\frac{1}{3}\imath k(\Phi + \Psi)\,, \tag{2.91}$$

$$\dot{\Theta}_\ell + \frac{\ell}{2\ell + 1}\imath k\Theta_{\ell-1} + \frac{\ell + 1}{2\ell + 1}\imath k\Theta_{\ell+1} = 0 \quad (\ell \geq 2)\,. \tag{2.92}$$

Gradients of the Bardeen potentials act as a source for the first moment. Because of the recursion relation, each multipole moment ℓ is coupled to the preceding and the following moment. Therefore, power is transferred to higher moments, and in principle we need to solve an infinite number of coupled differential equations. Simply truncating the hierarchy by imposing $\Theta_{\ell max} = 0$ is not an optimal solution, since the error due to the truncation will reflect back to lower moments via the coupling. A more effective truncation scheme is discussed in Ma & Bertschinger (1995). We notice that at early times and super-horizon scales (i.e. $k\eta \ll 1$) higher moments are suppressed by successive powers of $k\eta$, $\Theta_\ell \sim \mathcal{O}(\Theta_{\ell-1}k\eta)$, and hence the first few moments are sufficient to accurately describe the temperature fluctuation.

Relations with macroscopic quantities

From the definition of the stress-energy tensor (de Groot et al., 1980)

$$T^{\mu\nu}(x^\alpha) = \int \frac{\mathrm{d}^3 p}{p^0} p^\mu p^\nu f(x^\alpha, p^\mu) \tag{2.93}$$

[1] Different normalizations for the expansion coefficient are commonly used in the literature and their relation with the one used here is: in Hu & Sugiyama (1995b) $\Theta^{\mathrm{HS}} = \imath^\ell(2\ell + 1)\Theta_\ell$ (notice that in this work the Bardeen potentials are such that $\Psi^{\mathrm{HS}} = \Psi$ but $\Phi^{\mathrm{HS}} = -\Phi$); in Ma & Bertschinger (1995) Θ is denoted by Ψ and $\Psi_\ell^{\mathrm{MB}} = \imath^\ell\Theta_\ell$, which is the same convention used by Seljak & Zaldarriaga (1996); in Durrer (1994) Θ is denoted by $\mathcal{M}$ and $\mathcal{M}_\ell = \Theta_\ell/2$.

and comparing with Eq. (2.14, page 37), we can establish the hydrodynamical gauge invariant variables as integrals over momenta of the gauge invariant brightness perturbation:

$$D_{g,\gamma} = \frac{1}{\bar{\rho}_\gamma}\int \frac{\mathrm{d}\Omega}{4\pi}\mathcal{I}\,, \tag{2.94a}$$

$$V_\gamma^j = -\frac{1}{(1+w_\gamma)\bar{\rho}_\gamma}\int \frac{\mathrm{d}\Omega}{4\pi}n^j\mathcal{I}\,, \tag{2.94b}$$

$$\Pi^{ij} = \frac{1}{w_\gamma\bar{\rho}_\gamma}\int \frac{\mathrm{d}\Omega}{4\pi}n^{ij}\mathcal{I}\,. \tag{2.94c}$$

Rewriting the above in terms of multipole moments of the temperature perturbation, we have the identities in harmonic space[2]

$$\Theta_0 = \frac{1}{4}D_{g,\gamma}\,, \tag{2.95a}$$

$$\Theta_1 = -\frac{1}{3}\imath k V_\gamma\,, \tag{2.95b}$$

$$\Theta_2 = -\frac{1}{12}k^2\Pi_\gamma\,. \tag{2.95c}$$

Truncating the Boltzmann hierarchy at the third moment by setting $\Theta_\ell = 0$ for $\ell \geq 3$, we obtain

$$\dot{D}_{g,\gamma} + \frac{4}{3}k^2 V_\gamma = 0\,, \tag{2.96}$$

$$\dot{V}_\gamma - \frac{1}{4}D_{g,\gamma} = -\frac{1}{6}k^2\Pi_\gamma + \Phi + \Psi\,, \tag{2.97}$$

$$\dot{\Pi}_\gamma - \frac{8}{5}V_\gamma = 0\,. \tag{2.98}$$

Unsurprisingly, we recover the two conservation equations of (2.60-2.61, page 45) for radiation (with $w_\gamma = c_\gamma^2 = 1/3$ and $\Gamma = 0$), supplemented with an evolution equation for Π_γ. These equations are appropriate for relativistic, collisionless and massless particles such as neutrinos. At later times, however, higher order moments need to be taken into account. Photons are scattered by electrons, and to describe their evolution we now turn to the appropriate collision term.

2.2.5 Thomson scattering

We now consider the case of elastic Thomson scattering between photons and non-relativistic electrons. We give some elements of the derivation for the collision term for the total photon intensity, while we just outline the polarization treatment. A detailed derivation can be found in Kosowsky (1996); Durrer (2001).

Thomson scattering of unpolarized light generates linear polarization if the incident intensity has a quadrupolar anisotropy. In the tight coupling regime, collisions make the photons

[2]Notice that the monopole of our $\mathcal{F}$ corresponds (up to multiplicative constants) to the density perturbation in the comoving gauge; in the literature the temperature perturbation in Newtonian gauge is often employed (as in (Hu & Sugiyama, 1995b)), in which case an extra term $\propto \Phi$ appears along with Θ_0^{N}. With the normalization convention of (Hu & Sugiyama, 1995b), the relation between our monopole and the one in Newtonian gauge is $\Theta_0 = \Theta_0^{\mathrm{N}} - \Phi$. All other multipoles $\ell > 0$ do not suffer from this ambiguity and are gauge independent.

distribution function uniform in the electrons rest frame, and therefore no polarization can arise. However, during the weak coupling regime just before last scattering, the mean free path of photons grows and a sizable temperature quadrupole is generated, which acts as a source for polarization, as we briefly describe in this section. After decoupling, free streaming conserves the polarization state, which can only be changed by further rescattering due to reionization, see § 4.1.3.2.

2.2.5.1 Stokes parameters

The polarization state of light is usually described in terms of Stokes parameters, see e.g. Jackson (1975). The electric field of a plane monochromatic electromagnetic wave propagating in the z direction is

$$\mathbf{E}(\mathbf{x}, t) = \mathcal{E} e^{\imath(\omega t - kz)} \,, \tag{2.99}$$

where the complex vector $\mathcal{E}$ describing the polarization state of the wave is given by

$$\mathcal{E} = \begin{pmatrix} a_x e^{\imath\theta_x} \\ a_y e^{\imath\theta_y} \\ 0 \end{pmatrix} \,. \tag{2.100}$$

Instead of using the four numbers $(a_x, a_y, \theta_x, \theta_y)$, it is convenient to introduce the *Stokes parameters*

$$I \equiv a_x^2 + a_y^2 \,, \tag{2.101}$$
$$Q \equiv a_x^2 - a_y^2 \,, \tag{2.102}$$
$$U \equiv 2 a_x a_y \cos(\theta_x - \theta_y) \,, \tag{2.103}$$
$$V \equiv 2 a_x a_y \sin(\theta_x - \theta_y) \,, \tag{2.104}$$

which can be directly measured with a linear polarizer and a quarter-wave plate. Their physical interpretation is straightforward: I gives the total intensity, Q measures the difference between x and y polarization, U gives phase information for the two linear polarizations, and V determines the difference between positive and negative circular polarization. I and V are physical observables independent of the coordinate system, but Q and U mix under a rotation by an angle ϕ of the $x - y$ plane:

$$Q' = Q \cos(2\phi) + U \sin(2\phi) \tag{2.105a}$$
$$U' = -Q \sin(2\phi) + U \cos(2\phi) \,, \tag{2.105b}$$

from which it is easy to derive that the physically observable quantity is the polarization vector $\mathbf{P}$, lying in the $x - y$ plane, with magnitude $(Q^2 + U^2)^{1/2}$ and with polar angle $\alpha = \frac{1}{2} \tan^{-1} \frac{U}{Q}$.

Finally, the four stokes parameters are not independent, but satisfy the relation

$$I^2 = Q^2 + U^2 + V^2 \,. \tag{2.106}$$

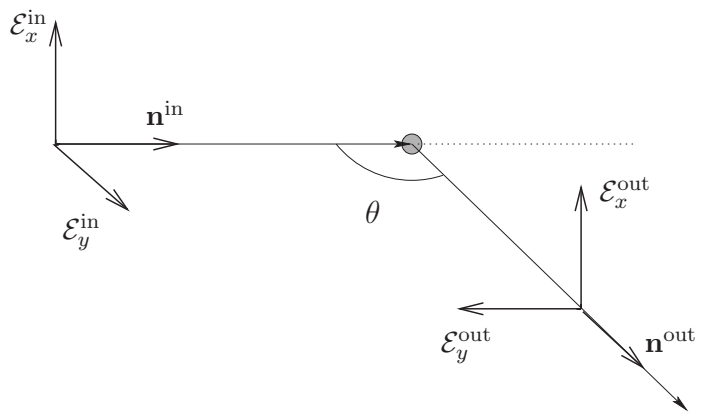

Figure 2.1: Geometry of the the Thomson scattering process in the rest frame of the electron, represented by the sphere in the center. A photon beam is incoming from the left and is scattered off with an angle θ.

2.2.5.2 Scattering cross section

We now consider the scattering process in the rest frame of the electron, with the geometry of Fig. 2.1. The Thomson scattering cross section for an incident wave with linear polarization $\mathcal{E}^{\rm in}$ into a scattered wave with polarization $\mathcal{E}^{\rm out}$ is

$$\frac{{\rm d}\sigma}{{\rm d}\Omega} = \frac{3\sigma_T}{8\pi}|\mathcal{E}^{\rm in}\cdot\mathcal{E}^{\rm out}|^2\,, \tag{2.107}$$

with σ_T the Thomson scattering cross section. It is convenient to work with the partial intensities I_x and I_y, defined as

$$I_x \equiv \frac{I+Q}{2} \quad \text{and} \quad I_y \equiv \frac{I-Q}{2}\,. \tag{2.108}$$

The incoming wave is unpolarized by assumption, so $I_x^{\rm in} = I_y^{\rm in} = I^{\rm in}/2$, and for the outgoing wave we find

$$I_x^{\rm out} = \frac{3\sigma_T}{16\pi} I^{\rm in} \quad \text{and} \quad I_y^{\rm out} = \frac{3\sigma_T}{16\pi} I^{\rm in}\cos^2(\theta) \tag{2.109}$$

or, in terms of the outgoing Stokes parameters

$$I^{\rm out} = \frac{3\sigma_T}{16\pi} I^{\rm in}(1+\cos^2(\theta))\,, \tag{2.110}$$

$$Q^{\rm out} = \frac{3\sigma_T}{16\pi} I^{\rm in}\sin^2(\theta)\,, \tag{2.111}$$

$$U^{\rm out} = 0\,. \tag{2.112}$$

The value of $U^{\rm out}$ has been found by recalculating Q in an outgoing basis which has been rotated by $\pi/4$. Thomson scattering does not generate circular polarization, so $V = 0$ and

we will not consider it further. Since from (2.106) there are only three independent Stokes parameters, and $V = 0$ all the time, the description in terms of I and Q is sufficient, and we wont use U any further.

The total outgoing intensities are obtained by integrating over all incoming directions, and rotating the result into a common coordinate system using (2.105):

$$I^{\rm out} = \frac{3\sigma_T}{16\pi}\int {\rm d}\Omega(1+\cos^2(\theta))I^{\rm in}(\theta,\phi)\,, \tag{2.113}$$

$$Q^{\rm out} = \frac{3\sigma_T}{16\pi}\int {\rm d}\Omega\sin^2(\theta)\cos(2\phi)I^{\rm in}(\theta,\phi)\,. \tag{2.114}$$

2.2.5.3 Temperature hierarchy

We are now in the position of deriving the collision term due to Thomson scattering for the intensity distribution function f, which is of the form

$$C\left[f\right] = \frac{{\rm d}f^+}{{\rm d}\eta} - \frac{{\rm d}f^-}{{\rm d}\eta}\,. \tag{2.115}$$

where $f^+(x^\mu, p^\mu)$ (f^-) denotes the distribution of particles within $(\Delta x^\mu, \Delta p^i/p^0)$ of (x^μ, p^μ) gained (lost) in the scattering process. According to the hypothesis of molecular chaos (de Groot et al., 1980), the contribution lost is just proportional to the electron density times the photon distribution, hence with the definitions (1.40–1.41, page 28)

$$\frac{{\rm d}f^-}{{\rm d}\eta}(x^\mu, p, n^i) = \dot\tau f(x^\mu, p, n^i)\,, \tag{2.116}$$

where

$$\dot\tau \equiv a\sigma_T n_e \tag{2.117}$$

is the *differential Thomson optical depth*, and n_e is the free electron density. The contribution scattered into $p^i = pn^i$ is most easily evaluated in the electron's rest frame, which we denote by a tilde. After averaging over incoming and summing over outgoing polarization states, we obtain

$$\frac{{\rm d}\tilde f^+}{{\rm d}\tilde t}(x^\mu, \tilde p, \tilde{\mathbf{n}}) = \sigma_T n_e \int \frac{{\rm d}\tilde\Omega_{\boldsymbol\varepsilon}}{4\pi}\tilde f(\tilde p, \tilde{\mathbf{n}})\omega(\tilde{\mathbf{n}}, \boldsymbol\varepsilon)\,, \tag{2.118}$$

where the angular dependence of the scattered intensity is, from (2.113)

$$\omega(\boldsymbol\varepsilon, \boldsymbol\varepsilon') = \frac{3}{4}[1-(\boldsymbol\varepsilon\cdot\boldsymbol\varepsilon')^2] = 1 + \frac{3}{4}\varepsilon_{ij}\varepsilon'^{ij} \tag{2.119}$$

with $\varepsilon_{ij} \equiv \varepsilon_i\varepsilon_j - \frac{1}{3}\delta_{ij}$. We now transform into the coordinate system, in which the photon distribution function f is defined. To first order we have the relations

$$\tilde p = p\left(1 + n_i(v_b^i - B^i)\right)\,, \tag{2.120}$$

$$\tilde{\mathbf{n}} = \mathbf{n}\,, \tag{2.121}$$

since aberration appears only at second order. We have used the baryon 3-velocity v_b^i, since electrons and baryons are electromagnetically coupled and their velocities are the same. Note that the above transformation assumes $v_b \ll 1$, i.e. that the electrons are non-relativistic, consistent with the fact that we consider v_b as a perturbation. Splitting the distribution

function in an isotropic part and a (gauge dependent) perturbation, $f = \bar{f}(\eta, p) + \delta f(x^i, p^i)$, we then compute the energy integrated collision term

$$4\pi \int p^3 \mathrm{d}pC\,[f] = a\sigma_T n_e \left[-4n_i(v_b^i - B^i)\bar{\rho}_\gamma + \delta\rho_\gamma - \delta I(\mathbf{n}) + \frac{3}{4} n^{ij}\delta I_{ij} \right] , \tag{2.122}$$

and we have introduced the gauge dependent brightness perturbation $\delta I \equiv 4\pi \int \mathrm{d}p p^3 \delta f$ and its second moment

$$\delta I_{ij} \equiv \int \frac{\mathrm{d}\Omega_{\boldsymbol{\varepsilon}}}{4\pi} \varepsilon_{ij} \delta I(\boldsymbol{\varepsilon})\,. \tag{2.123}$$

The expression Eq. (2.122) can be brought in explicit gauge invariant form by substituting the gauge dependent variables with the corresponding gauge independent counterparts. After some manipulations we obtain

$$4\pi \int p^3 \mathrm{d}pC\,[f] = 4\dot{\tau}\bar{\rho}_\gamma \left[\Theta_0 - n_i V_b^i - \Theta + \frac{1}{16} n_{ij} \Pi_\gamma^{ij} \right] , \tag{2.124}$$

where we have used the identity (2.94c, page 49). In view of adding the collision term on the right hand side of the hierarchy (2.90, page 48), it is convenient to rewrite it in terms of multipoles of the temperature fluctuation Θ and transform to Fourier space

$$4\pi \int p^3 \mathrm{d}pC\,[f] = -4\dot{\tau}\bar{\rho}_\gamma \left[(\imath k V_b + 3\Theta_1) P_1 + \frac{9}{2}\Theta_2 P_2 + \sum_{\ell \geq 3} (2\ell + 1)\Theta_\ell P_\ell \right] . \tag{2.125}$$

A few remarks are in order at this point: as a consequence of the conservation of energy in the elastic collision, non-relativistic Thomson scattering does not contain a monopole, while the dipole corresponds to a velocity mismatch between photons and baryons, as is apparent from the first term on the right hand side with $3\Theta_1 = -\imath k V_\gamma$. The angular dependence of the scattering generates a quadrupole moment. In the limit of very many collisions, $\dot{\tau} \gg \mathcal{H}$, all multipoles $\ell > 1$ are driven to zero, therefore in the *strong coupling regime*, the photons and baryons velocity coincide and higher order moments are suppressed: thus the tight-coupled photons-baryons system can be described as an hydrodynamical fluid in term of the zeroth and first moments only.

The Boltzmann hierarchy, Eq. (2.90, page 48), supplemented with the above collision term for photons-electrons Thomson scattering, now becomes:

$$\dot{\Theta}_0 + \imath k \Theta_1 = 0\,, \tag{2.126a}$$

$$\dot{\Theta}_1 + \frac{1}{3}\imath k(\Theta_0 + \Phi + \Psi) + \frac{2}{3}\imath k \Theta_2 = -\dot{\tau}\left(\frac{1}{3}\imath k V_b + \Theta_1\right) , \tag{2.126b}$$

$$\dot{\Theta}_2 + \frac{2}{5}\imath k \Theta_1 + \frac{3}{5}\imath k \Theta_3 = -\dot{\tau}\frac{9}{10}\Theta_2 \tag{2.126c}$$

$$\dot{\Theta}_\ell + \frac{\ell}{2\ell+1}\imath k \Theta_{\ell-1} + \frac{\ell+1}{2\ell+1}\imath k \Theta_{\ell+1} = -\dot{\tau}\Theta_\ell \quad (\ell \geq 3)\,. \tag{2.126d}$$

Rewriting the above in terms of macroscopic quantities and cutting the hierarchy at $\ell = 2$ gives instead of Eq. (2.96, page 49)

$$\dot{D}_{g,\gamma} + \frac{4}{3}k^2 V_\gamma = 0\,, \tag{2.127a}$$

$$\dot{V}_\gamma - \frac{1}{4}D_{g,\gamma} + \frac{1}{6}k^2 \Pi_\gamma - \Phi - \Psi = -\dot{\tau}(V_\gamma - V_b)\,, \tag{2.127b}$$

$$\dot{\Pi}_\gamma - \frac{8}{5}V_\gamma = -\dot{\tau}\frac{9}{10}\Pi_\gamma\,. \tag{2.127c}$$

2.2.5.4 Polarization hierarchy

As discussed in § 2.2.5.2, photons scattered at a right angle are are preferentially polarized along the direction orthogonal to the scattering plane (i.e. in the $\mathcal{E}_x^{\rm out}$ direction in Fig. 2.1 when $\theta = \pi/2$). Expanding the incoming intensity in spherical harmonics according to

$$I^{\rm in}(\theta,\phi) = \sum_\ell \sum_m I_{\ell m} Y_{\ell m}(\theta,\phi)\,, \tag{2.128}$$

then the resulting $Q^{\rm out}$, from (2.114) is

$$Q^{\rm out} = \frac{3\sigma_T}{4\pi}\sqrt{\frac{2\pi}{15}}{\rm Re}\ I_{22}\,, \tag{2.129}$$

which shows that if the incoming photon intensity as a function of direction has a non-zero component of Y_{22}, associated with an $\ell = 2$ quadrupolar moment, then there will be a net linear polarization of the outgoing distribution.

In analogy with the intensity distribution function f, we denote by $f^Q = \bar{f}^Q(\eta,p) + F^Q(\eta,x^i,p,n^i)$ the perturbed distribution function in phase space and by Θ^Q the brightness perturbation for the Stokes parameter Q,

$$\Theta^Q = \frac{1}{4}\frac{\int_0^\infty \bar{f}^Q p^3 {\rm d}p}{\int_0^\infty F^Q p^3 {\rm d}p}\,. \tag{2.130}$$

Then the collisional Boltzmann equation for the brightness perturbation f^Q in Fourier space is (Bond & Efstathiou, 1984; Kosowsky, 1996)

$$\dot{\Theta}^Q + \imath k\mu\Theta^Q = -\dot{\tau}\left[\Theta^Q + \frac{1}{2}(1-P_2)\left(\Theta_2 + \Theta_2^Q - \Theta_0^Q\right)\right]\,. \tag{2.131}$$

Expanding the equation in Legendre polynomials as in Eq. (2.85, page 48), we obtain the Boltzmann polarization hierarchy:

$$\dot{\Theta}_0^Q + \imath k\Theta_1^Q = -\frac{\dot{\tau}}{2}\left[\Theta_2 + \Theta_0^Q + \Theta_2^Q\right]\,, \tag{2.132}$$

$$\dot{\Theta}_1^Q + \frac{1}{3}\imath k\left[\Theta_0^Q + 2\Theta_2^Q\right] = -\dot{\tau}\Theta_1^Q\,, \tag{2.133}$$

$$\dot{\Theta}_2^Q + \frac{2}{5}\imath k\Theta_1^Q + \frac{3}{5}\imath k\Theta_3^Q = -\frac{\dot{\tau}}{10}\left[9\Theta_2^Q - \Theta_2 + \Theta_0^Q\right]\,, \tag{2.134}$$

$$\dot{\Theta}_\ell^Q + \frac{\ell}{2\ell+1}\imath k\Theta_{\ell-1}^Q + \frac{\ell+1}{2\ell+1}\imath k\Theta_{\ell+1}^Q = -\dot{\tau}\Theta_\ell^Q \quad (\ell \geq 3)\,. \tag{2.135}$$

Polarization effects also feed back into the temperature collision term, modifying the $\ell = 2$ equation in the temperature hierarchy (2.126) as follows:

$$\dot{\Theta}_2 + \frac{2}{5}\imath k\Theta_1 + \frac{3}{5}\imath k\Theta_3 = -\frac{\dot{\tau}}{10}\left[9\Theta_2^Q - \Theta_2 + \Theta_0^Q\right]\,. \tag{2.136}$$

2.2.5.5 E and B polarization

From the the hierarchy of equations (2.132) it is possible to determine the brightness perturbation for Q today, and define the corresponding power spectrum. However, the approach using Stokes parameters is limited by the fact that U and Q are not rotationally invariant,

but are defined with respect to a fixed coordinate system on the sky. Not only the superposition of different modes is cumbersome because of the behavior of Q and U under rotation, but the coordinate system becomes ambiguous and ill-defined on the whole sky, since it is impossible to define a rotationally invariant orthogonal basis on the two-sphere.

The solution is to construct two spin 2 quantities from Q and U, which one then expands in the appropriate spin-weighted basis on the two-sphere (Zaldarriaga & Seljak, 1997), and reduces to scalar quantities by acting on them with spin raising and lowering operators. This manipulations yield two scalar quantities which are rotationally invariant, and therefore well defined on the whole sky. Furthermore, one can expand these quantities in terms of usual spherical harmonics and build two linear combinations which behave differently under parity transformation: the combination labelled E, in analogy with the electric field, is invariant under a parity change, while the B-type combination changes it sign, analogous to the magnetic field. Another terminology, sometimes found in the literature, is C mode for "curl" (corresponding to the B-type) and G for "gradient" (corresponding to the E-type).

Another advantage of this decomposition is that only the cross-correlation between E-polarization and temperature is needed, since the cross-correlation between B and E or T vanishes since B has opposite parity. Furthermore, scalar modes do not generate B polarization, due to the peculiar μ dependence of Thomson scattering, while tensor modes do. Therefore, the separation of the polarization signal in E and B modes is useful to separate scalar from tensor contribution, and to identify foreground contamination or a lensing signal, which can convert E polarization into B polarization for scalar modes.

We do not give explicit expressions here, which are rather technical and are not needed in the following, but refer the reader to Zaldarriaga & Seljak (1997) instead. A similar decomposition, but with a different normalization has been proposed by Kamionkowski et al. (1997).

Nous partons des faits, pour composer des théories, et nous tâchons toujours de nous éloigner le moins possible de ces faits. Nous ignorons ce qu'est *l'essence* des choses, et n'en avons cure, parce qu'une telle étude sort de notre domaine.

Vilfredo Pareto
Traité de sociologie générale

Part II

COSMIC MICROWAVE BACKGROUND

Chapter 3

Fundamental equations

The all sky picture of CMB anisotropy delivered by COBE and more recently and with 30 times more resolution by WMAP can be considered as a fingerprint of the early Universe. More precisely, it is an accurate reproduction of the fluctuations in the radiation-matter mixture at the epoch of recombination.

In this section we succinctly explain the origin of this picture, by starting with the behavior of scalar perturbations in a Universe containing one perfect fluid, § 3.1; many of the fundamental features of the anisotropies can be understood in a simple model with a mixture of radiation and matter which are coupled only gravitationally, as demonstrated in § 3.2 where the concepts of adiabatic and CDM isocurvature initial conditions are introduced; adding a massless neutrino component yields two new growing modes, the neutrino entropy/density and velocity isocurvature solutions, derived in § 3.3. Although the results of those two sections are already known in the literature, the derivation presented in this work is original. We then refine the picture of acoustic oscillations by including baryons in § 3.4, and sketch the origin of damping in § 3.5. Finally we derive the line of sight solution for the observed temperature fluctuations today and introduce the CMB angular power spectra in § 3.6. The understanding and tools developed in the following will build the basis for the next chapters, where parameter extraction techniques will be discussed (Chapter 5) and applications presented (Chapters 6 and 7).

There is a rich literature on the cosmic microwave background but unfortunately an updated work which encompasses both and introduction to the field and more advanced material, covering the rapid evolution of the last few years, is presently lacking. Throughout this and the next chapter we give ample references to the classic and more recent research papers; as background material, Lineweaver et al. (1997) is a valuable source which presents an introduction to the CMB theory as well as some observational issues; Durrer (2001) is built on a gauge invariant formalism similar to the one used here; Partridge (1995) is a good introductory overview written at the onset of the recent data-driven epoch. A rather complete review of both theory and data analysis is offered by Hu & Dodelson (2002).

3.1 One perfect fluid

We begin by examining the behavior of scalar perturbations in a flat ($\mathcal{K} = 0$) universe which contains a single perfect fluid, described by $w = c_s^2 = \text{const}$, and $\Gamma = \Pi = 0$.

Since the anisotropic stress vanishes, from Eq. (2.51, page 44) it follows $\Psi = \Phi$. The evolution of the perturbations is given by the two conservation equations (2.62–2.63, page 45) supplemented by the Poisson equation (2.49, page 44), which in Fourier space read:

$$\dot{D} - 3w\mathcal{H}D = -(1+w)k^2V\,, \tag{3.1}$$

$$\dot{V} + \mathcal{H}V = \Psi + \frac{c_s^2}{1+w}D\,, \tag{3.2}$$

$$-k^2\Psi = \frac{3}{2}\mathcal{H}^2D\,. \tag{3.3}$$

These equations can be combined into a second order equations for the density contrast:

$$\ddot{D} + (1-3w)\mathcal{H}\dot{D} - \frac{3}{2}\mathcal{H}^2(1+2w-3w^2)D + c_s^2k^2D = 0 \tag{3.4}$$

By defining a new variable $x \equiv k\eta$ and the parameter $\nu \equiv 2/(1+3w)$, we obtain the following equation for $\mathcal{D} \equiv Dx^{\nu-2}$

$$\frac{\mathrm{d}^2}{\mathrm{d}x^2}\mathcal{D} + \frac{2}{x}\frac{\mathrm{d}}{\mathrm{d}x}\mathcal{D} + \left[c_s^2 - \frac{\nu(\nu+1)}{x^2}\right]\frac{\mathcal{D}}{x^2} = 0\,, \tag{3.5}$$

For $c_s^2 \neq 0$ the solution is a linear combination of spherical bessel (j_ν) and von Neumann (n_ν) functions of order ν (Abramowitz & Stegun, 1970)

$$\mathcal{D} = C_1 j_\nu(c_s x) + C_2 n_\nu(c_s x) \equiv Z_\nu(c_s x)\,. \tag{3.6}$$

Therefore the general solution of Eqs. (3.1) is

$$D = x^{2-\nu}Z_\nu(c_s x)\,, \tag{3.7}$$

$$V = \frac{3}{2}\nu\left[Z_\nu(c_s x)x^{1-\nu} + \frac{2-\nu}{3\nu(1+\nu)}x^{2-\nu}Z_{\nu-1}(c_s x)\right]\,, \tag{3.8}$$

$$\Psi = -\frac{3}{2}\nu^2 x^{-\nu}Z_\nu(c_s x)\,. \tag{3.9}$$

The asymptotic behavior of the Bessel and von Neumann functions is

$$j_\nu \propto x^\nu \quad \text{for } c_s x \ll 1, \qquad j_\nu \propto \frac{1}{x}\cos(c_s x - \gamma_\nu) \quad \text{for } c_s x \gg 1, \tag{3.10a}$$

$$n_\nu \propto x^{-(\nu+1)} \quad \text{for } c_s x \ll 1, \qquad n_\nu \propto \frac{1}{x}\sin(c_s x - \gamma_\nu) \quad \text{for } c_s x \gg 1. \tag{3.10b}$$

with $\gamma_\nu \equiv \pi(\nu+1)/2$. For an expanding universe ($x > 0$) and $\nu > -1$ (i.e. $w < -1$ or $w > -1/3$) n_ν is divergent at early times, $c_s x \ll 1$. Therefore we set $C_2 = 0$ and we obtain the asymptotic solutions (for $w > -1/3$)

$$\begin{cases} \Psi = \Psi_0 \\ D = -\dfrac{2}{3}\dfrac{\Psi_0}{\nu^2}x^2 \\ kV = \dfrac{2}{(1+\nu)\nu^2}\Psi_0 x \end{cases} \qquad \text{for } c_s x \ll 1 \tag{3.11}$$

and

$$\begin{cases} \Psi = \Psi_0 x^{-(1+\nu)} \cos(c_s x + \gamma_\nu) \\ D = -\dfrac{2}{3}\dfrac{\Psi_0}{\nu^2} x^{1-\nu} \cos(c_s x + \gamma_\nu) \\ kV = \dfrac{(\nu-2)\Psi_0}{3(1+\nu)} x^{1-\nu} \cos(c_s x + \gamma_{\nu-1}) \end{cases} \qquad \text{for } c_s x \gg 1. \tag{3.12}$$

This solution was first discovered by Bardeen (1980). The Bardeen potential is constant on super-horizon scales, and decays once inside the acoustic horizon. On scales smaller than the acoustic horizon ($c_s x \ll 1$) density perturbations oscillate: the gravitational attraction is resisted by the fluid pressure ($w \neq 0$) and this sets up acoustic oscillations. The amplitude of density and velocity fluctuations remains constant inside the horizon in the case of radiation ($\nu = 1, w = 1/3$), while it increases for $w > 1/3$ or $w > -1/3$. The behavior of the density and velocity perturbations on scales larger then the horizon depends on the variable under consideration. While D, corresponding to the density contrast in the comoving gauge, is growing, the density contrast in the flat slicing gauge D_g remains constant. Therefore there is no universal criterion to establish the growth of perturbations outside the horizon: the behavior depends on the chosen gauge. As we go to early times, $x \to 0$, perturbation theory remains valid as long as it is possible to find a gauge in which the largest perturbation variable does not diverge. We come back to this point in § 4.3, where we derive the most general initial conditions.

The case of dust $w = c_s^2 = 0$ has a power-law solution on all scales. It suffices to remark that Eq. (3.5) reduces to

$$\frac{\mathrm{d}^2}{\mathrm{d}x^2} D + \frac{2}{x}\frac{\mathrm{d}}{\mathrm{d}x} D - \frac{6}{x^2} D = 0\,, \tag{3.13}$$

whose general solution is $D = Ax^2 + Bx^{-3}$. The growing exact solution is therefore

$$\begin{cases} \Psi = \Psi_0 \\ D = -\dfrac{1}{6}\Psi_0 x^2 \propto a \\ kV = \dfrac{1}{3}\Psi_0 x \propto a^{1/2} \end{cases} \qquad \text{for dust, } w = 0. \tag{3.14}$$

Clearly, in a dust universe perturbations always grow on sub-horizon scales, since there is no pressure to counterbalance the gravitational attraction.

3.2 Cold dark matter and radiation

In this section we investigate the evolution of perturbations in a flat universe containing only radiation and a pressureless matter component which is decoupled from radiation. Thus the matter has only a gravitational effect and represents a cold dark matter component. In the next section we include massless decoupled neutrinos in the picture, while the role of baryons, which are coupled to photons via Thomson scattering, is investigated in § 3.4.

3.2.1 Adiabatic and isocurvature modes

In this section we use as density variable the density contrast in the total comoving gauge Δ_α, defined in Eq. (2.37, page 42). We identify the radiation with photons (subscript γ),

and we have $w_\gamma = c_\gamma^2 = 1/3$, while for matter $w_m = c_m^2 = 0$. We normalize the scale factor at the matter-radiation equality, so that

$$\bar{\rho}_m(a_{\rm eq}) = \bar{\rho}_\gamma(a_{\rm eq}) \quad \text{with} \quad a_{\rm eq} \equiv 1 \quad \text{hence} \quad \frac{\bar{\rho}_m}{\bar{\rho}_\gamma} = a\,. \tag{3.15}$$

The total equation of state parameter and sound velocity are therefore

$$w = \frac{1}{3}\frac{1}{a+1} \quad \text{and} \quad c_s^2 = \frac{1}{3}\frac{4}{4+3a}\,. \tag{3.16}$$

As long as we are considering times well before decoupling, the photons form a tight coupled fluid with baryons, since Thomson scattering prevents the generation of anisotropic stress (and higher multipoles in the Boltzmann hierarchy) in the photons component, $\Pi_\gamma = 0$, as we show in § 3.4. Therefore, via the anisotropic stress equation (2.51, page 44), the Bardeen potentials are equal, $\Psi = \Phi$. The Bardeen equation for Φ (2.67, page 46) is then

$$\ddot{\Phi} + 3\mathcal{H}(1+c_s^2)\dot{\Phi} + 3(c_s^2 - w)\mathcal{H}^2\Phi = c_s^2\triangle\Phi + \tfrac{3}{2}\mathcal{H}^2 w\Gamma\,, \tag{3.17}$$

where $\Gamma = \Gamma_{\rm rel}$ is related to the relative entropy perturbation $S \equiv S_{m\gamma} = \Delta_m - \frac{3}{4}\Delta_\gamma$ by Eq. (2.41, page 42). By using the Poisson equation we can rewrite the above as an equation for the total density contrast,

$$\begin{aligned}\mathcal{H}^{-2}\ddot{D} + (1 - 6w + 3c_s^2)\mathcal{H}^{-1}\dot{D} - \tfrac{3}{2}(1 + 8w - 3w^2 - 6c_s^2)D = \\ - c_s^2\left(\frac{k}{\mathcal{H}}\right)^2\left[D - 3c_z^2(1+w)S\right]\,,\end{aligned} \tag{3.18}$$

where we have introduced $c_z^2 \equiv \bar{\rho}_\gamma\bar{\rho}_m(c_\gamma^2 - c_m^2)/\left[(1+w)\bar{\rho}\right] = a/(3a+4)$.

The energy conservation equation (2.60, page 45) reads for the radiation and matter components:

$$\dot{D}_{g,\gamma} + \frac{4}{3}k^2 V_\gamma = 0 \quad \text{(radiation)}, \tag{3.19}$$

$$\dot{D}_{g,m} + k^2 V_c = 0 \qquad \text{(matter)}. \tag{3.20}$$

Subtracting (3.20) from (3.19) and using that

$$\frac{D_{g,\alpha}}{1+w_\alpha} - \frac{D_{g,\beta}}{1+w_\beta} = \frac{\Delta_\alpha}{1+w_\alpha} - \frac{\Delta_\beta}{1+w_\beta} = S_{\alpha,\beta} \tag{3.21}$$

we obtain

$$\dot{S} = -k^2(V_m - V_\gamma)\,. \tag{3.22}$$

In order to find an evolution equation for the entropy S, we derive (3.22) and making use of the momentum conservation equation (2.63, page 45) after a lengthy manipulation we arrive at

$$\mathcal{H}^{-2}\ddot{S} + (1 - 3c_z^2)\mathcal{H}^{-1}\dot{S} = \left(\frac{k}{\mathcal{H}}\right)^2\left[\frac{1}{3(1+w)}D - c_z^2 S\right]\,. \tag{3.23}$$

Together, Eqs. (3.18) and (3.23) describe the evolution of adiabatic (curvature) and isocurvature (dark matter) perturbations in a flat universe containing only dark matter and radiation.

We start by considering large scales ($k \ll \mathcal{H}$) at early times, $a \to 0$. Then the right hand side of (3.18) and (3.23) is negligible, thus D and S are decoupled. Using the scale factor a as variable, we obtain an homogeneous system

$$\begin{cases} a^2 \dfrac{\mathrm{d}^2}{\mathrm{d}a^2} D - 2D = 0 \\ a^2 \dfrac{\mathrm{d}^2}{\mathrm{d}a^2} S + a \dfrac{\mathrm{d}}{\mathrm{d}a} S = 0 \end{cases} \qquad \text{(Large scales, radiation epoch)} \tag{3.24}$$

whose general solution consists of four modes,

$$\begin{cases} D = D_0 a^2 + D_1 a^{-1} \\ S = S_0 + S_1 \ln a \end{cases} . \tag{3.25}$$

We will call the mode with $D_0 \neq 0, D_1 = S_0 = S_1 = 0$ the *growing adiabatic mode*, while the one with $S_0 \neq 0, D_0 = D_1 = S_1$ the *growing isocurvature mode* (notice that for $a < 1$ the S_1 mode is indeed decaying). As we show below, the isocurvature mode at early times has vanishing total density contrast, Bardeen potential curvature perturbation, $\zeta = 0$, hence its name[1].

Consider first the growing adiabatic mode: we can now restore the solution for D in the source term on the right hand side of Eq. (3.23) to find the solution for S up to second order in $k/\mathcal{H}$. The Bardeen potential is easily found from the Poisson equation, and the result is

$$\begin{cases} D = D_0 a^2 \\ S = \dfrac{D_0}{64} \left(\dfrac{k}{\mathcal{H}} \right)^2 a^2 \propto a^4 \\ \Phi = -\dfrac{3D_0}{2} \left(\dfrac{\mathcal{H}a}{k} \right)^2 = \text{ const} \\ kV = \dfrac{1}{2} \dfrac{k}{\mathcal{H}} \Phi \propto a \\ \zeta = -\dfrac{9D_0}{4} \left(\dfrac{\mathcal{H}a}{k} \right)^2 = \text{ const} \end{cases} \qquad \text{(adiabatic, radiation epoch).} \tag{3.26}$$

Clearly, we recover the behavior already found in the single radiation fluid case for the potential. We also discover that the entropy perturbation grows as a^4, but remains negligible on large scales, thus the adiabaticity condition $S \approx 0$ is maintained on large scales.

For the growing isocurvature mode we find, to the same approximation

$$\begin{cases} D = \dfrac{S_0}{12} \left(\dfrac{k}{\mathcal{H}} \right)^2 a \propto a^3 \\ S = S_0 \\ \Phi = -\dfrac{S_0}{8} a \\ kV = -\dfrac{S_0}{8} \dfrac{k}{\mathcal{H}} a \propto a^2 \\ \zeta = -\dfrac{3S_0}{16} a \end{cases} \qquad \text{(isocurvature, radiation epoch).} \tag{3.27}$$

[1]The CDM isocurvature mode is sometimes termed "isothermal" in the literature: this comes from the fact that $D = 0$ implies $\frac{\delta T}{T} = -\frac{\bar{\rho}_m}{\bar{\rho}_\gamma} \Delta_m \approx 0$ at early times. Intuitively, it takes only a small perturbation in the radiation component to compensate for a fluctuation in the matter at early times, because the Universe is radiation dominated.

We see that there is no generation of entropy on large scales ($\dot{S} = 0$), however the isocurvature condition $\Phi \approx 0$ is maintained only as long as $a \ll 1$. Naively we would expect that, as long as the scale considered is outside the horizon, the term containing S on the right hand side of Eq. (3.18) is suppressed as $k^2/\mathcal{H}^2$, thus D (hence Φ) should not grow significantly. However, since $\Phi \propto \mathcal{H}^2/k^2$, effects of magnitude $k^2/\mathcal{H}^2$ in D are significant for Φ. This can be seen more directly by rewriting the right hand side of Eq. (3.17) as $-c_s^2 k^2/\mathcal{H}^2\Phi - 2(1+w)c_s^2 c_z^2 S$. Therefore even on super-horizon scale the term $\propto S$ act as a source for Φ whenever $c_s^2 c_z^2$ is significantly non-zero. This is the case during the transition from the radiation to the matter dominated epoch.

Having established the behavior in the early epoch, we now turn our attention to scales which enter the horizon when the universe is well matter dominated, i.e. to wavelengths such that

$$k \ll k_{\rm eq} \equiv \mathcal{H}(a_{\rm eq})\,. \tag{3.28}$$

The effects of the radiation-matter transition are easiest to discuss by looking at the behavior of the curvature perturbation ζ. To this end we rewrite the evolution equation (2.69, *page*46) as

$$\dot{\zeta} = -c_s^2 \mathcal{H}\left[\frac{2}{3(1+w)}\left(\frac{k}{\mathcal{H}}\right)^2 \Phi + 3c_z^2 S\right]\,. \tag{3.29}$$

The term $\propto \Phi$ on the right hand side is always negligible on super-horizon scales ($k/\mathcal{H} \ll 1$); for adiabatic perturbations we also have $S = 0$, and thus we obtain

$$\zeta = \text{ const} \qquad \text{(adiabatic, all times)}, \tag{3.30}$$

the usual conservation law for ζ in the adiabatic case. For the isocurvature mode ($S = S_0 =$ const) we find by integration

$$\zeta = -3S_0 \int_0^a \frac{{\rm d}a}{a} c_s^2 c_z^2 \underset{a\to\infty}{\longrightarrow} -\frac{1}{3}S_0 \quad \text{(isocurvature, matter epoch)}. \tag{3.31}$$

The radiation-matter transition generates a curvature perturbation from the initial isocurvature one, and this even on super-horizon scales.

Since $\zeta = $ const in the matter era independently on the initial conditions, we can find the value of the Bardeen potential in the matter epoch simply by integrating the definition of the curvature perturbation, using that $w =$ const as well. We then obtain the relation (valid only in the regime where $\zeta =$ const, $w =$ const)

$$\Phi = \frac{3(1+w)}{5+3w}\zeta + Ca^{-\frac{5+3w}{2}}\,, \tag{3.32}$$

and we can drop the second term, which is decaying for $w > -5/3$. Therefore

$$\Phi(a \gg a_{\rm eq}) = \text{ const } = \tfrac{3}{5}\zeta \quad \text{(matter epoch, independent of IC)}. \tag{3.33}$$

For the adiabatic mode, $\zeta =$ const in the radiation era as well, therefore we can apply (3.32) with $w \approx$ const $= 1/3$, getting

$$\Phi(a \ll a_{\rm eq}) = \text{ const } = \tfrac{2}{3}\zeta \quad \text{(radiation epoch, adiabatic)}. \tag{3.34}$$

Let us denote by Φ_0 the value of Φ at the moment when the initial conditions for the perturbations are specified, deep in the radiation era. The adiabatic mode corresponds to $S_0 = 0, \Phi_0 \neq 0$, while the isocurvature mode has $S_0 \neq 0, \Phi_0 = 0$. From (3.33) we know that Φ is constant on super-horizon scales in the matter era, independent of the type of initial conditions; we denote its value by $\Phi_{\rm MD}$, and we wish to express it in terms of S_0, Φ_0. For adiabatic perturbations, ζ stays constant through the transition, and therefore combining (3.33) with (3.34)

$$\Phi_{\rm MD} \approx \tfrac{9}{10}\Phi_0 \quad \text{(adiabatic, large scales)}. \tag{3.35}$$

For isocurvature perturbations, the growth of ζ through the transition gives a non-zero Φ in the matter epoch, from (3.33) and (3.31) :

$$\Phi_{\rm MD} \approx -\frac{1}{5}S_0 \quad \text{(isocurvature, large scales)}. \tag{3.36}$$

In conclusion, we can summarize our results in terms of a transfer matrix as

$$\begin{pmatrix} \Phi \\ S \end{pmatrix}_{a \gg a_{\rm eq}} = \begin{pmatrix} 9/10 & -1/5 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \Phi_0 \\ S_0 \end{pmatrix} . \tag{3.37}$$

It is often useful to use the curvature perturbation as a variable describing the adiabatic mode, instead of Φ. In terms of the initial values of the curvature and entropy perturbations, (ζ_0, S_0), the final values in the matter era are given by a transfer matrix of the form

$$\begin{pmatrix} \zeta \\ S \end{pmatrix}_{a \gg a_{\rm eq}} = \begin{pmatrix} T_{\zeta\zeta} & T_{\zeta S} \\ 0 & T_{SS} \end{pmatrix} \begin{pmatrix} \zeta_0 \\ S_0 \end{pmatrix} . \tag{3.38}$$

From the above analysis, we conclude that for scales $k \ll k_{\rm eq}$ the transfer coefficients are

$$T_{\zeta\zeta} = 1\,, \quad T_{\zeta S} = -\frac{1}{3}\,, \quad T_{SS} = 1\,. \tag{3.39}$$

For smaller scales, which enter the horizon before the universe is completely matter dominated, the coefficients have to be found numerically.

3.2.2 Acoustic oscillations

We have seen in § 3.1 that perturbations in a fluid of photons oscillate on scales smaller than the horizon. We now discuss the corresponding behavior in the presence of matter, and link the phase of the oscillations to the adiabatic or isocurvature initial conditions on large scales.

Neglecting the anisotropic stress, $\Pi_\gamma = 0$, the conservation equations (2.60–2.61, page 45) for photons read

$$\dot{D}_{g,\gamma} + \tfrac{4}{3}k^2 V_\gamma = 0 \tag{3.40}$$

$$\dot{V}_\gamma - \tfrac{1}{4}D_{g,\gamma} = 2\Phi \tag{3.41}$$

where Φ can be considered as an external potential determined by the Poisson equation. We can recast the above in a second order equation for the density perturbation:

$$\ddot{D}_{g,\gamma} + c_\gamma^2 k^2 D_{g,\gamma} = 2\Phi\,. \tag{3.42}$$

Adiabatic initial conditions

Let's consider Eq. (3.42) deep in the matter era, when the driving force is just a constant set by the dominating matter contribution in the adiabatic case. Then the general solution of Eq. (3.42) is

$$D_{g,\gamma} = C_1 \cos(c_\gamma k\eta) + C_2 \sin(c_\gamma k\eta) - 8\Phi \tag{3.43}$$

$$kV_\gamma = \frac{1}{4c_\gamma}\left[C_1 \sin(c_\gamma k\eta) - C_2 \cos(c_\gamma k\eta)\right] . \tag{3.44}$$

For small scales, where all choices of density perturbation are equivalent, we recover the oscillatory behavior already found in § 3.1. The density perturbations perform harmonic oscillations around a zero point displaced by a constant factor.

The constants C_1 and C_2 are fixed by the initial conditions, adiabatic or isocurvature, established by matching the above solution on large scales with the results of the previous section. To this end, we shall use the following relation between $D_{g,\gamma}$ and Δ_γ, which follows from the definitions of the variables:

$$\tfrac{1}{4}D_{g,\gamma} = \tfrac{1}{3}\Delta_m - \tfrac{1}{3}S - \mathcal{H}V - \Phi\,. \tag{3.45}$$

From the momentum conservation equation (2.63, page 45) we obtain for the total velocity perturbation in the matter era

$$\dot{V} + \mathcal{H}V = \Phi\,, \tag{3.46}$$

with solution

$$V = V_1 a^{-1} + \frac{2}{3}\mathcal{H}^{-1}\Phi\,. \tag{3.47}$$

The term $\propto a^{-1}$ is decaying, therefore we retain $V \sim \frac{2}{3}\mathcal{H}^{-1}\Phi$. Inserting this into Eq. (3.45) and using that in the matter era $\Phi = 9/10\Phi_0 - S_0/5$ we obtain on large scales, where $\Delta_m \sim (k/\mathcal{H})^2\Phi \ll \Phi$,

$$\tfrac{1}{4}D_{g,\gamma}(a \gg a_{\rm eq}) \approx \text{const } = -\tfrac{3}{2}\Phi_0\,. \tag{3.48}$$

Thus on large scales and in the matter epoch, $D_{g,\gamma}$ is independent of the entropy perturbation, and is simply related to the primordial Bardeen potential.

The adiabatic mode stays decoupled from the isocurvature mode on super-horizon scales, therefore we can set the initial conditions for the solution (3.43–3.44) by taking its constant-time super-horizon limit, i.e. $k \to 0$, $\eta = \text{const} \gg \eta_{\rm eq}$. This gives, with $S_0 = 0$

$$\tfrac{1}{4}D_{g,\gamma} = \tfrac{1}{4}C_1 - 2\Phi_{\rm MD} \tag{3.49}$$

and comparing with Eq. (3.48) and using again (3.35) we obtain

$$C_1 = \tfrac{4}{3}\Phi_{\rm MD}\,. \tag{3.50}$$

The constant C_2 is set by noting that the adiabatic condition $S = 0$ is preserved on super-horizon scales, and that, because of energy-momentum conservation for matter and radiation, this implies

$$V_\gamma = V_m\,. \tag{3.51}$$

Since

$$V = \frac{4}{4+3a}V_\gamma + \frac{3a}{4+3a}V_m \tag{3.52}$$

we have that $V \approx V_m$ for $a \gg a_{\rm eq}$, and with (3.47) it follows that

$$V_\gamma = V_m \approx \tfrac{2}{3}\Phi\mathcal{H}^{-1}\,. \tag{3.53}$$

Comparing this with the large scale limit of Eq. (3.44),

$$\lim_{k\to 0,\eta=\text{const}} V_\gamma = \frac{\eta}{4}\left[C_1 - C_2 \lim_{y\to 0}\frac{\cos y}{y}\right]\,, \tag{3.54}$$

we see that we need to impose $C_2 = 0$, otherwise V_γ would diverge in the large-scale limit $y \to 0$, and we recover again (3.50) by using $\mathcal{H} = 2/\eta$:

$$C_1 = \tfrac{4}{3}\Phi_{\rm MD} \quad \text{and} \quad C_2 = 0\,. \tag{3.55}$$

In conclusion, the adiabatic solution is

$$\begin{cases} D_{g,\gamma} = \dfrac{4}{3}\Phi\cos(c_\gamma k\eta) - 8\Phi \\ kV_\gamma = c_\gamma\Phi\sin(c_\gamma k\eta) \end{cases} \qquad \text{(adiabatic)}. \tag{3.56}$$

Isocurvature initial conditions

As we have seen in the previous section, $\Phi = 0$ is no longer maintained in the matter era for isocurvature initial conditions. It is therefore convenient to solve (3.42) at early times in the radiation regime, where we know that the driving term on the right hand side is $\Phi \propto \eta$ (cf. Eq. (3.27, page 63)):

$$D_{g,\gamma} = C_1\cos(c_\gamma k\eta) + C_2\sin(c_\gamma k\eta) - \frac{3}{4}k^{-2}\eta_{\rm eq}^{-1}S_0\eta\,, \tag{3.57}$$

$$kV_\gamma = \frac{1}{4c_\gamma}\left[C_1\sin(c_\gamma k\eta) - C_2\cos(c_\gamma k\eta)\right] + \frac{9}{16}k^{-3}\eta_{\rm eq}^{-1}S_0\,. \tag{3.58}$$

The constants C_1 and C_2 are determined by looking at the early time limit on super-horizon scales, $\eta \to 0$, $k = \text{const} \ll k_{\rm eq}$. From the early-times solution (3.27) we have that $D_{g,\gamma} \to 0$ for $\eta \to 0$, and therefore we need to set $C_1 = 0$. The early time limit for Eq. (3.58) gives

$$\lim_{\eta\to 0,k=\text{const}} kV_\gamma = -\frac{C_2}{4c_\gamma} + \frac{9}{16}k^{-3}\eta_{\rm eq}^{-1}S_0\,, \tag{3.59}$$

while from the isocurvature solution (3.27) combined with (3.52) we have for $a \ll a_{\rm eq}$

$$\lim_{\eta\to 0,k=\text{const}} kV_\gamma = kV \propto \eta^2 \to 0\,. \tag{3.60}$$

By requiring that the left hand side of (3.59) vanishes we conclude that

$$C_2 = \frac{3}{4c_\gamma}k^{-3}\eta_{\rm eq}^{-1}S_0\,. \tag{3.61}$$

In conclusion, isocurvature initial conditions excite a sine oscillation in the radiation density:

$$\begin{cases} D_{g,\gamma} = \dfrac{3}{4}k^{-2}\eta_{\rm eq}^{-1}S_0\left[\sqrt{3}k\sin(c_\gamma k\eta) - \eta\right] \\ kV_\gamma = -\dfrac{3\sqrt{3}}{16}k^{-2}\eta_{\rm eq}^{-1}S_0\left[\sqrt{3}k\cos(c_\gamma k\eta) - 1\right] \end{cases} \qquad \text{(isocurvature)}. \tag{3.62}$$

An heuristic argument (Hu & Sugiyama, 1995b) explains why adiabatic initial conditions excite the cosine mode while isocurvature initial conditions produce the sine mode: at early times, the potential acting as a driving force on the right hand side of Eq. (3.42) is constant for adiabatic initial conditions, while it is $\propto \eta$ in the isocurvature case. This mimics a cosine and a sine forcing term, respectively, and thus the corresponding modes get excited. An approximated analytical solution valid until recombination and through the radiation-matter transition can be found in Hu & Sugiyama (1995a).

3.3 Neutrinos and initial conditions

In this section we extend the above treatment to include massless neutrinos. They are described as an additional relativistic component, which is decoupled from the others below a temperature of a few MeV, and therefore their distribution function obeys the collisionless Boltzmann equation. We shall see in the following that the anisotropic stress created by free streaming of neutrinos considerably complicates the simple picture of the previous section.

By including one more component in the mixture, we generally expect two additional modes to arise, which we will be able to identify with the so-called "neutrino isocurvature density" (NID) and "neutrino isocurvature velocity" (NIV) modes. In the following, we shall refer to both of them as to "neutrino isocurvature modes"[2], and we will sometimes call the neutrino density mode "neutrino entropy", which is a more appropriate definition in our view. These two modes were first found by Bucher et al. (2000), who solved a formal expansion in powers of η of the Einstein and conservation equations at early times and on large scales (i.e. for $\eta k \to 0$) in synchronous gauge, an analysis repeated in the gauge invariant formalism in Trotta (2001). The approach we propose here offers a more physical understanding and the approximations we employ could be extended to a refined analytical model of the sub-horizon structure of the neutrino modes angular power spectra. We explicitly give some details of the derivation, since to our knowledge this calculation is new.

We argue in § 3.3.4 that an "anisotropic stress mode", which is characterized by a non-vanishing Π_ν at early times, is non-physical, since it leads to incurable divergences in the perturbation variables.

3.3.1 Evolution equations for a three components model

In the presence of neutrinos, the background radiation energy density is written as

$$\bar{\rho}_r = \bar{\rho}_\gamma + \bar{\rho}_\nu = \bar{\rho}_\gamma (1 + r_\nu)\,, \tag{3.63}$$

where we have defined the constant $r_\nu \equiv (7N_\nu/8)(4/11)^{4/3} \approx 0.68$ for $N_\nu = 3$ neutrino families. As before, the scale factor is normalized to matter-radiation equality, the conformal Hubble parameter is

$$\mathcal{H} = \frac{1 + \eta/2}{\eta + \eta^2/4} = \frac{(1 + 7a)^{1/2}}{7a}\,, \tag{3.64}$$

[2]The term "isocurvature" is somewhat abused for the neutrino density mode, see the remark after Eq. (3.93) on page 73. We nevertheless employ this terminology for simplicity and consistency with the literature.

and the cosmological parameters as a function of the scale factor are of the form

$$\Omega_\nu(a) = \frac{r_\nu}{(1+r_\nu)(1+a)}\,, \tag{3.65}$$

$$\Omega_\gamma(a) = \frac{1}{(1+r_\nu)(1+a)}\,, \tag{3.66}$$

$$\Omega_m(a) = \frac{a}{(1+a)}\,. \tag{3.67}$$

We still neglect the dynamical effect of baryons, which to lowest order is unimportant, but continue to assume that Thomson scattering drives to zero all multipoles $\ell \geq 2$ in the Boltzmann hierarchy for photons, which are then described as a relativistic perfect fluid. Neutrinos become collisionless after neutrino decoupling, therefore the fluid approximation is insufficient. A neutrino anisotropic stress is generated by free streaming and to lowest order we cut the Boltzmann hierarchy for neutrinos, Eq. (2.96, page 49), by setting to zero all moments ≥ 3. The goal is to derive second order evolution equations for the three relevant and physical quantities: the total density contrast D, the entropy perturbations in the dark matter, $S_{m\gamma}$, and in the neutrinos, $S_{\nu\gamma}$, supplemented by an evolution equation for the neutrino anisotropic stress.

The source term in the Bardeen equation is modified in two ways: there is an additional entropy contribution coming from the neutrino entropy perturbation $S_{\nu\gamma}$, and we have to take into account the anisotropic stress term. This gives for the evolution equation of the total density contrast D (compare with (3.18, page 62))

$$\begin{aligned}\mathcal{H}^{-2}\ddot{D} + (1-6w+3c_s^2)\mathcal{H}^{-1}\dot{D} - \tfrac{3}{2}(1+8w-3w^2-6c_s^2)D = \\ -\left(\frac{k}{\mathcal{H}}\right)^2\left\{\left[c_s^2 D - 3c_s^2c_z^2(1+w)\left(S_{m\gamma} - \frac{r_\nu}{1+r_\nu}S_{\nu\gamma}\right)\right]\right. \\ \left. + \frac{2r_\nu}{3(1+r_\nu)(1+a)}\left[\mathcal{H}\dot{\Pi}_\nu - \left[(1+3w) - 3c_z^2\right]\mathcal{H}^2\Pi_\nu - \frac{1}{2}k^2\Pi_\nu\right]\right\}.\end{aligned} \tag{3.68}$$

Equation (3.23, page 62) acquires extra terms coming from $S_{\nu\gamma}$, reading

$$\mathcal{H}^{-2}\ddot{S}_{m\gamma} + (1-3c_z^2)\mathcal{H}^{-1}\left[\dot{S}_{m\gamma} - \dot{S}_{\nu\gamma}\right] = \left(\frac{k}{\mathcal{H}}\right)^2\left[\frac{1}{3(1+w)}D - c_z^2 S_{m\gamma} - \frac{4wr_\nu}{3(1+w)(1+r_\nu)}S_{\nu\gamma}\right]. \tag{3.69}$$

In deriving the above equations we have made use of (2.38, page 42) and (3.22, page 62) together with the following useful relations:

$$\frac{1}{4}\Delta_\gamma = \frac{1}{3(1+w)}D - \frac{4r_\nu}{3(4+3a)(1+r_\nu)}S_{\nu\gamma} - \frac{a}{4+3a}S_{m\gamma}\,, \tag{3.70}$$

$$kV_\gamma = kV - \frac{4r_\nu}{(4+3a)(1+r_\nu)}(V_\nu - V_\gamma) - \frac{3a}{4+3a}(V_m - V_\gamma)\,. \tag{3.71}$$

We obtain an equation for the neutrino entropy perturbation by deriving the difference of the momentum conservation equation for neutrinos (Eq. (2.97, page 49) written for ν instead of γ) and the momentum conservation for the photon fluid, (3.19, page 62), with the result

$$\ddot{S}_{\nu\gamma} + \frac{k^2}{3}S_{\nu\gamma} = \frac{k^4}{6}\Pi_\nu\,. \tag{3.72}$$

The coupled system (3.68), (3.69) and (3.72) describes the evolution of adiabatic and entropy perturbations in a mixture of photons, dark matter and radiation, once we specify Π_ν. However, on super-horizon scales and for early times, $k/\mathcal{H} \ll 1$, the anisotropic stress is unimportant, since from (2.98, page 49) written for ν instead than for γ, it obeys

$$a\frac{\mathrm{d}}{\mathrm{d}a}k^2\Pi_\nu = \frac{8}{5}\frac{k}{\mathcal{H}}kV_\nu \approx 0\,, \tag{3.73}$$

which shows that on super-horizon scales there is no generation of anisotropic stress, a result expected on the grounds of causality arguments. At earlier times, the neutrinos were coupled to electrons via weak interaction processes, which isotropized the neutrino distribution function suppressing any appreciable anisotropic stress; hence we can assume that at the time under consideration (just after neutrino decoupling) there is no anisotropic stress to zeroth order in powers of a, i.e. $\Pi_\nu = \mathcal{O}(a)$ at least.

In the above approximation and for $a \ll 1$ we thus obtain the simple system

$$\begin{cases} a^2\dfrac{\mathrm{d}^2}{\mathrm{d}a^2}D - 2D = 0\,, \\ a^2\dfrac{\mathrm{d}^2}{\mathrm{d}a^2}S_{m\gamma} + a\dfrac{\mathrm{d}}{\mathrm{d}a}S_{m\gamma} = a\dfrac{\mathrm{d}}{\mathrm{d}a}S_{\nu\gamma}\,, \\ a^2\dfrac{\mathrm{d}^2}{\mathrm{d}a^2}S_{\nu\gamma} = 0\,, \end{cases} \tag{3.74}$$

whose general solution consists of six modes,

$$\begin{cases} D = D_0 a^2 + D_1 a^{-1}\,, \\ S_{m\gamma} = S_0 + S_1 \ln a + N_v a\,, \\ S_{\nu\gamma} = N_d + N_v a\,. \end{cases} \tag{3.75}$$

We recognize the growing and decaying adiabatic (the D_0 and D_1 terms, respectively) and isocurvature dark matter (S_0 and S_1 terms, respectively) modes, and we also find two new non-decaying modes, a constant neutrino entropy mode N_d, and a neutrino velocity mode $N_v a$ (the reason for this terminology is explained below).

In order to go beyond this large scales solution, we need to include the effect of the anisotropic stress. To this end, we recast Eq. (3.73) by substituting kV_ν with

$$kV_\nu = kV - \frac{a\mathcal{H}}{k}\left[\frac{1}{1+r_\nu}\frac{\mathrm{d}}{\mathrm{d}a}S_{\nu\gamma} + \frac{3(1+r_\nu)a}{4}\frac{\mathrm{d}}{\mathrm{d}a}S_{m\gamma}\right]\,. \tag{3.76}$$

From now on we drop the last term on the right hand side, which is always suppressed by a power of a except in the dark matter isocurvature case, which we do not investigate further here. For the total velocity, the constraint equation (2.50, page 44), combined with the the anisotropic stress equation (2.51, page 44) and the Poisson equation (2.49, page 44) yield, in the early time $a \ll 1$ limit

$$kV = \frac{\mathcal{H}}{k}\left(\frac{3}{4}D - \frac{3a}{4}\frac{\mathrm{d}}{\mathrm{d}a}D - \frac{r_\nu}{1+r_\nu}k^2\Pi_\nu\right)\,. \tag{3.77}$$

The evolution equation (3.73) for the anisotropic stress then reads, for $a \ll 1$

$$a\frac{\mathrm{d}}{\mathrm{d}a}k^2\Pi_\nu + \frac{4}{5}\frac{r_\nu}{1+r_\nu}k^2\Pi_\nu = \frac{6}{5}D - \frac{6a}{5}\frac{\mathrm{d}}{\mathrm{d}a}D - \frac{8a}{5(1+r_\nu)}\frac{\mathrm{d}}{\mathrm{d}a}S_{\nu\gamma}\,. \tag{3.78}$$

In the same limit and in terms of the scale factor a, the equations for D and $S_{\nu\gamma}$ become (dropping the last term $\propto k^2\Pi_\nu$ on the right hand side of (3.68) which is always negligible compared to the others):

$$a^2\frac{\mathrm{d}^2}{\mathrm{d}^2 a}D - 2D = -\left(\frac{k}{\mathcal{H}}\right)^2 \frac{r_\nu}{3(1+r_\nu)}S_{\nu\gamma} - \frac{2r_\nu}{3(1+r_\nu)}\left[a\frac{\mathrm{d}}{\mathrm{d}a}k^2\Pi_\nu - 2k^2\Pi_\nu\right], \tag{3.79}$$

$$a^2\frac{\mathrm{d}^2}{\mathrm{d}^2 a}S_{\nu\gamma} + \frac{1}{3}\left(\frac{k}{\mathcal{H}}\right)^2 S_{\nu\gamma} = \frac{1}{6}\left(\frac{k}{\mathcal{H}}\right)^2 k^2\Pi_\nu\,. \tag{3.80}$$

The system of coupled differential equations (3.78), (3.79) and (3.80) is too difficult to solve analytically. To find an approximate solution valid to leading order in powers of a for early times, we treat the anisotropic stress iteratively as a perturbation to the large scale solution, Eq. (3.75), in analogy with the procedure in Hu & Sugiyama (1995a). More specifically, we use the large scale solution for D and $S_{\nu\gamma}$ as a source on the right hand side of Eq. (3.78) to determine the anisotropic stress, then we re-insert the solution for Π_ν on the right hand side of (3.79) and (3.80) to find self-consistent corrections to the large scale behavior.

As an illustration, let us first consider the adiabatic growing mode, $D = D_0 a^2, D_1 = S_0 = S_1 = N_d = N_v = 0$. In that case, the right hand side of (3.78) is dominated by the terms in D, giving

$$a\frac{\mathrm{d}}{\mathrm{d}a}k^2\Pi_\nu + \frac{4}{5}\frac{r_\nu}{1+r_\nu}k^2\Pi_\nu = -\frac{6}{5}D_0 a^2\,, \tag{3.81}$$

which has the particular solution

$$k^2\Pi_\nu = -\frac{3(1+r_\nu)D_0}{7r_\nu+5}a^2\,. \tag{3.82}$$

Notice that, although the above form of $\Pi_\nu \propto a^2$ is of the same order as the adiabatic solution $D \propto a^2$, its contribution on the right hand side of (3.79) cancels out because of the factor 2 in the exponent. Thus it is consistent to have neglected the anisotropic stress in the first place when deriving the large scale solution.

With the above approximation for Π_ν, from (3.80) we can determine the growth of neutrino entropy perturbations in the adiabatic mode, finding to leading order in powers of a

$$S_{\nu\gamma} = -\frac{(1+r_\nu)D_0}{48(7r_\nu+5)}\left(\frac{k}{\mathcal{H}}\right)^2 a^2 \propto a^4 \ll D\,. \tag{3.83}$$

The growth of the dark matter entropy perturbation is also modified by the coupling to the neutrino entropy perturbations on the left hand side of (3.69, page 69), but the term $\propto \dot{S}_{\nu\gamma} \propto a^4$ has the same scaling as the term $\propto D$ on the right hand side, and the approximate solution is

$$S_{m\gamma} = \frac{1}{64}\left[1 - \frac{1+r_\nu}{3(7r_\nu+5)}\right]D_0\left(\frac{k}{\mathcal{H}}\right)^2 a^2 \propto a^4 \ll D\,. \tag{3.84}$$

In conclusion, the growing adiabatic mode at early times in the presence of neutrinos and

anisotropic stress has the approximate solution (compare with the solution (3.26, page 63)):

$$\begin{cases} D = D_0 a^2 \\ S_{m\gamma} \propto \left(\dfrac{k}{\mathcal{H}}\right)^2 a^2 \propto a^4 \\ S_{\nu\gamma} \propto \left(\dfrac{k}{\mathcal{H}}\right)^2 a^2 \propto a^4 \\ k^2\Pi_\nu \propto a^2 \\ \Phi = -\dfrac{3}{2}\left(\dfrac{ka}{\mathcal{H}}\right)^2 D_0 = \Phi_0 = \text{ const} \qquad \text{(adiabatic).} \\ \Psi = \Phi_0 + \dfrac{3r_\nu}{7(1+r_\nu)}\left(\dfrac{ka}{\mathcal{H}}\right)^2 \equiv \Psi_0 = \text{ const} \\ kV = \dfrac{1}{2}\dfrac{k}{\mathcal{H}}\Phi_0 \propto a \\ \zeta = -\dfrac{9D_0}{4}\left(\dfrac{\mathcal{H}a}{k}\right)^2 = \text{ const} \end{cases} \tag{3.85}$$

The Bardeen potentials are no longer equal due to the anisotropic stress, the fractional correction being

$$\left|\frac{\Phi_0 - \Psi_0}{\Phi_0}\right| = \frac{2}{7}\frac{r_\nu}{1+r_\nu} \approx 0.1\,, \tag{3.86}$$

of order 10%, in good agreement with Hu & Sugiyama (1995a).

3.3.2 Neutrino entropy mode

Let us now turn our attention to the $N_d \neq 0$ mode, with $N_v = D_0 = D_1 = S_0 = S_1 = 0$: this is clearly a neutrino entropy mode, since $S_{\nu\gamma} = \text{const}$ for $a \to 0$.

To determine the growth of perturbations in the total density D beyond the large scale solution $D = 0$, consider the right hand side of Eq. (3.79): if the anisotropic stress goes at least as a^2, then the part containing Π_ν cancels (for $\Pi_\nu \propto a^2$) or is subdominant with respect to the $S_{\nu\gamma}$ term (for $\Pi_\nu = \mathcal{O}(a^3)$ or higher). In any case, we can neglect the anisotropic stress term as a source for D with respect to the neutrino entropy perturbation, with the caveat that at the end of our calculation we have to check that this assumption is satisfied - indeed, cf. Eq. (3.89). By this argument, we look for a particular solution of

$$a^2\frac{\mathrm{d}^2}{\mathrm{d}^2 a}D - 2D = -\left(\frac{k}{\mathcal{H}}\right)^2 \frac{r_\nu}{3(1+r_\nu)}N_d\,, \tag{3.87}$$

which is given by

$$D = -\frac{r_\nu}{9(1+r_\nu)}N_d\left(\frac{k}{\mathcal{H}}\right)^2 \ln(a) \propto a^2\ln(a)\,. \tag{3.88}$$

The logarithmic dependence can be neglected if we do not apply this solution over a too large time range (say, less than a few orders of magnitude), and replaced by the value of $\ln(a)$ evaluated at the typical value of the scale factor in the range considered, a_*, which we reabsorb in the overall normalization by defining a new constant $N_d^* \equiv N_d \ln(a_*)$.

We can now solve for Π_ν by inserting the above expression for D in Eq. (3.78, page 70), and observing that on the right hand side $\frac{\mathrm{d}S_{\nu\gamma}}{\mathrm{d}a} = 0$, thus obtaining

$$k^2\Pi_\nu = N_d^* \left(\frac{k}{\mathcal{H}}\right)^2 \frac{r_\nu}{3(7r_\nu+5)} \propto a^2\,, \tag{3.89}$$

which is consistent with our initial assumption for Π_ν.

Finally, the Bardeen potentials follow from the Poisson equation and the anisotropic stress equation, yielding

$$\Phi = \frac{r_\nu N_d^*}{6(1+r_\nu)} = \text{ const}\,, \tag{3.90}$$

$$\Psi = \Phi\left(1 - \frac{2r_\nu}{7r_\nu+5}\right) = \text{ const}\,. \tag{3.91}$$

The gauge invariant curvature perturbation ζ is given by (2.27, page 40) and it can be rewritten as

$$\zeta = \frac{3}{2}\Phi + \frac{a}{2}\frac{\mathrm{d}}{\mathrm{d}a}\Phi - \frac{r_\nu}{2(1+r_\nu)}\left(\frac{\mathcal{H}}{k}\right)^2 k^2\Pi_\nu\,. \tag{3.92}$$

yielding for the neutrino entropy mode

$$\zeta = \frac{r_\nu N_d^*}{1+r_\nu}\left(\frac{1}{4} - \frac{r_\nu}{6(7r_\nu+5)}\right) = \text{ const}\,. \tag{3.93}$$

This results agree with the power law solution found by Bucher et al. (2000), which they called "neutrino isocurvature density" mode; we prefer however to term this mode "neutrino entropy", since the initial curvature perturbation does not vanish, and indeed is of the same order as the entropy perturbation.

3.3.3 Neutrino velocity mode

The mode with $N_v \neq 0$ has vanishing entropy at early times, since $S_{\nu\gamma} \to 0$ for $a \to 0$, but the bulk velocity difference between neutrinos and photons in non-zero,

$$k(V_\nu - V_\gamma) = -\frac{\dot{S}_{\nu\gamma}}{k} = \text{const} \tag{3.94}$$

hence its name.

From the power-law solution for this mode (see Bucher et al., 2000; Trotta, 2001) we expect that the anisotropic stress goes to leading order as $\Pi_\nu \propto a$. Indeed, by replacing the large-scale solution $D = 0, S_{\nu\gamma} = N_v a$ on the right hand side of (3.78) we find the particular solution

$$k^2\Pi_\nu = -\frac{8N_v}{9r_\nu+5}a\,. \tag{3.95}$$

We now use this expression as a source on the right hand side of (3.79) to determine the corrections to D, and we can ignore the contribution of the term $\propto S_{\nu\gamma}$ which goes as a^3 compared to the part containing Π_ν, which is dominant, being proportional to a. We thus have to solve

$$a^2\frac{\mathrm{d}^2}{\mathrm{d}a^2}D - 2D = -\frac{16r_\nu N_v}{3(1+r_\nu)(9r_\nu+5)}a\,, \tag{3.96}$$

and we find the particular solution

$$D = \frac{8r_\nu N_v}{3(1+r_\nu)(9r_\nu+5)}a\,. \tag{3.97}$$

As already noticed in Bucher et al. (2000), the Bardeen potentials are decaying

$$\Phi = -\frac{4r_\nu N_v}{(1+r_\nu)(9r_\nu+5)}\left(\frac{\mathcal{H}}{k}\right)^2 a \propto a^{-1}\,, \tag{3.98}$$

$$\Psi = -\Phi\,, \tag{3.99}$$

but this does not necessarily mean that perturbation theory breaks down for $a \to 0$. In general, a solution is considered non divergent if it is possible to find a gauge in which *all* the perturbation variables do no diverge in the limit $a \to 0$. The synchronous gauge potentials for the neutrino velocity mode are indeed non-singular at early times (Bucher et al., 2000). In fact, even though the Bardeen potential diverge, the gauge invariant curvature perturbation ζ vanishes to leading order. This is most easily seen by making use of Eq. (2.32d, page 41), finding

$$\zeta = \frac{1}{2}\left(\Psi + \Phi\right) = 0\,, \tag{3.100}$$

and thus the velocity mode is indeed an isocurvature mode.

The leading order corrections to $S_{m\gamma} = 0$ induced by the neutrino modes can be obtained as particular solutions to Eq. (3.69, page 69), which for early times reads

$$a^2\frac{\mathrm{d}}{\mathrm{d}a^2}S_{m\gamma} + a\left[\frac{\mathrm{d}}{\mathrm{d}a}S_{m\gamma} - \frac{\mathrm{d}}{\mathrm{d}a}S_{\nu\gamma}\right] = -\left(\frac{k}{\mathcal{H}}\right)^2\frac{r_\nu}{3(1+r_\nu)}S_{\nu\gamma}\,. \tag{3.101}$$

Summarizing, the early time solutions for neutrino entropy ($N_d \neq 0$) and neutrino isocurvature velocity ($N_v \neq 0$) initial conditions are:

Neutrino entropy	**Neutrino velocity**
$S_{\nu\gamma} = N_d$	$S_{\nu\gamma} = N_v a$
$D = -\left(\frac{k}{\mathcal{H}}\right)^2\frac{r_\nu N_d^*}{9(1+r_\nu)} \propto a^2$	$D = \frac{8r_\nu N_v}{3(1+r_\nu)(9r_\nu+5)}a$
$S_{m\gamma} = -\left(\frac{k}{\mathcal{H}}\right)^2\frac{r_\nu N_d}{12(1+r_\nu)} \propto a^2$	$S_{m\gamma} = aN_v$
$kV = \frac{1}{2}\frac{k}{\mathcal{H}}\Psi \propto a$	$kV = \frac{k}{\mathcal{H}}\Psi = \text{ const}$
$k^2\Pi_\nu = \left(\frac{k}{\mathcal{H}}\right)^2\frac{r_\nu N_d^*}{3(7r_\nu+5)} \propto a^2$	$k^2\Pi_\nu = -\frac{8N_v}{9r_\nu+5}a$
$\Phi = \frac{r_\nu N_d^*}{6(1+r_\nu)} = \text{ const}$	$\Phi = -\frac{4r_\nu N_v}{(1+r_\nu)(9r_\nu+5)}\left(\frac{\mathcal{H}}{k}\right)^2 a \propto a^{-1}$
$\Psi = \Phi\left(1 - \frac{2r_\nu}{7r_\nu+5}\right) = \text{ const}$	$\Psi = -\Phi$
$\zeta = \frac{r_\nu N_d^*}{1+r_\nu}\left(\frac{1}{4} - \frac{r_\nu}{6(7r_\nu+5)}\right) = \text{ const}$	$\zeta = 0\,.$

(3.102)

3.3.4 The divergent nature of the anisotropic stress mode

One could ask whether it would be possible to excite a growing "neutrino anisotropic stress mode", characterized by initial conditions $D = S_{\nu\gamma} = S_{m\gamma} = V_{\nu\gamma} = V_{m\gamma} = 0$ and $\Pi_\nu \neq 0$ for $a \to 0$. Even though highly exotic, such a mode, if it existed, should be included if we want to consider the most general type of perturbations. We now show that this mode is divergent in all gauges, and therefore is non-physical, since it would lead to the breakdown of perturbation theory for $a \to 0$. Alternatively, we can see it as a decaying mode, which therefore does not need to be considered since it quickly disappears.

Consider the anisotropic stress equation (2.51, page 44) with $\Pi_\nu = \Pi_0 = \text{const}$ on the right hand side,

$$\Psi = \Phi - \frac{r_\nu}{(1+r_\nu)(1+a)}\mathcal{H}^2\Pi_0\,. \tag{3.103}$$

Since $\mathcal{H} = \eta^{-1}$ to leading order for $a \ll a_{\rm eq}$, it follows that $\Psi \propto \eta^{-2}$. The fact that the Bardeen potential diverges at early times is not by itself sufficient to discard the corresponding mode, as we have seen in the example of the neutrino velocity mode. A necessary condition, however, is the existence of a gauge in which all of the perturbation variables constructed out of $A, B, C, E, \delta, v, \pi_L$ are non-divergent. For the neutrino velocity mode, this gauge is the synchronous gauge. Clearly, since Ψ is a gauge invariant variable, by construction it does not change under a gauge transformation but the variables A, B, C, E do, according to the transformation laws (2.18, page 38). If we expand in a Laurent series around $\eta = 0$ the definition of Ψ, Eq. (2.25b, page 40), and we allow terms η^n with exponent $n \geq -2$, because of $\mathcal{H} = 1/\eta$ we obtain to leading order

$$A = \Psi \propto \eta^{-2}\,. \tag{3.104}$$

In other words, in the radiation dominated universe a metric perturbation of the form $A \propto \eta^{-2}$ is gauge invariant. This can also be seen directly from the transformation law for A, Eq. (2.18a, page 38): the part $\mathcal{H}T + \dot{T}$ does not contain terms $\propto \eta^{-2}$ if T is written as a Laurent series in η. We conclude that $\Pi_0 \neq 0$ induces a divergence of A for early times, which does not disappear in any gauge. One could conceive to combine A with other diverging variables to construct via cancellation a non-diverging metric variable: this however would unavoidably produce divergent terms in the matter variables. Therefore a neutrino anisotropic stress mode is always decaying in all gauges.

In principle, there is a whole hierarchy of modes coming from setting $\Theta^\ell_\nu \neq 0$ for $\ell \geq 3$ as initial conditions in the neutrino Boltzmann hierarchy. As we noticed in § 2.2.4, higher order moments are coupled to the potentials and to the velocity and density perturbations by successive powers of $k\eta$. By reversing the argument, we see that $\Theta^\nu_{\ell-1} = \mathcal{O}\left(\Theta^\nu_\ell/k\eta\right)$ implies that in the early Universe and on super-horizon scales, $k\eta \ll 1$, choosing $\Theta_\ell = \mathcal{O}(1)$ for $\ell \geq 3$ would produce divergent behavior in the lower-order multipoles of the hierarchy. Since for $\ell \geq 2$ the multipole moments are gauge invariant, it follows that there is no gauge in which such a mode is growing. In summary, the adiabatic and the general isocurvature modes presented above constitute the most general type of perturbation.

3.4 The role of baryons

In this section, we go back to the model of a Universe containing dark matter and photons, and refine the treatment given in § 3.2 by taking into account the role of baryons in the dynamic of the oscillations. For simplicity, we neglect the corrections induced by the neutrinos anisotropic stress, omitting neutrinos entirely.

Before recombination photons interact with electrons via Thomson scattering (see section 2.2.5). The time-scale for the scattering process is set by the Compton scattering time $\dot{\tau}^{-1}$, which represents the typical time between two collisions. Tight coupling is an expansion in powers of $\dot{\tau}^{-1}$, assuming that the scattering rate is rapid enough to equilibrate changes in the photon-baryons fluid, and in this limit moments $\ell \geq 2$ in the photon distribution function are suppressed by successive powers of $\dot{\tau}^{-1}$. Therefore to lowest order the photon distribution function is described by its zeroth and first multipoles only, and we can set $\Pi_\gamma = \Theta_{\ell\geq 3} = 0$, which justifies the approximation taken in the previous section. Therefore the truncated Boltzmann hierarchy (2.127, page 53) gives for photons

$$\dot{D}_{g,\gamma} + \frac{4}{3}k^2 V = 0\,, \tag{3.105}$$

$$\dot{V}_\gamma - \frac{1}{4}D_{g,\gamma} - 2\Phi = -a\sigma_T n_e (V_\gamma - V_b)\,. \tag{3.106}$$

To ensure conservation of the total momentum, we need to supplement the conservation equation for baryons with the *Thomson drag force* term coming from the scattering process, obtained as the first moment of the collision term

$$F_j^{\rm drag} = a\sigma_T n_e \rho_\gamma \int \frac{{\rm d}\Omega}{4\pi} n_j C\left[f\right]\,. \tag{3.107}$$

The momentum conservation for baryons, Eq. (2.60, page 45), therefore gives

$$\dot{D}_{g,b} + k^2 V_b = 0 \tag{3.108}$$

$$\dot{V}_b + \mathcal{H}V_b - \Phi = -\frac{1}{R}a\sigma_T n_e (V_b - V_\gamma)\,, \tag{3.109}$$

and we have defined $R \equiv 3\bar{\rho}_b/(4\bar{\rho}_\gamma)$, which can easily be estimated

$$R \approx \left(\frac{670}{1+z}\right)\left(\frac{\Omega_b h^2}{0.022}\right)\,. \tag{3.110}$$

The set of Eqs. (3.105–3.106) and (3.108–3.109) describes the evolution of perturbations for the tight-coupled photon-baryon fluid, while the dark matter component enters via its influence on the gravitational potential Φ. To lowest order in $1/\dot{\tau}$, collisions force the baryons and photons velocities to coincide, $V_\gamma = V_b$, which via Eq. (3.22, page 62) implies $\dot{S}_{b\gamma} = 0$, hence the entropy per baryon is conserved.

Equations (3.105, 3.106 and 3.109) can now be combined into the equation of a damped, forced harmonic oscillator:

$$\frac{\rm d}{{\rm d}\eta}\left[(1+R)\dot{D}_{g,\gamma}\right] + \frac{k^2}{3}D_{g,\gamma} = -\frac{4}{3}(2+R)k^2\Phi\,. \tag{3.111}$$

By comparing with Eq. (3.42, page 65), we see that baryons have two effects: they change the effective mass of the system (factor $(1+R)$ on the left hand side) and they displace the zero

point of the oscillation by adding to the potential Φ. Both modifications are a consequence of the fact that baryon add to the mass of the system but not to the restoring pressure, which is still given by the photons alone.

The time dependence of R is of the order of the Hubble time, hence large compared to the time scale of one oscillation. For illustrative purpose, we can then neglect the time dependence of R and obtain from Eq. (3.111)

$$\ddot{D}_{g,\gamma} + c_s^2 k^2 D_{g,\gamma} = -4(2+R)c_s^2 k^2 \Phi \,, \tag{3.112}$$

where the sound speed of the coupled fluid is $c_s^2 = 1/(3(1+R))$. At early times, $c_s^2 \to 1/3$, as appropriate for radiation, while at late times $c_s^2 \approx 0$, when the universe is dominated by matter. The homogeneous solution is still a superposition of sine and cosine oscillations, but adding the baryons slows down the period by decreasing c_s^2 with respect to the pure photons fluid. This is responsible for a shift in the acoustic peak positions and for a larger distance between the peaks in the CMB power spectrum, see the explanations regarding the role of the shift parameter on page 101.

The adiabatic solution (3.56) becomes

$$D_{g,\gamma} = \frac{4}{3}(1+R)\Phi\cos(c_s k\eta) - 4(2+R)\Phi \,, \tag{3.113}$$

$$kV_\gamma = \left(\frac{1+R}{3}\right)^{1/2} \Phi \sin(c_s k\eta) \,. \tag{3.114}$$

The amplitude of the cosine oscillation has increased by a factor $(1+R)$, and the potential well has deepened by an extra factor $(1+R/2)$. This displacement of the zero point of the oscillations induces a boost (decrease) of the odd (even) peaks in the power spectrum sometimes denotes as "baryon driving", which is discussed in § 4.1.2.2 and shown in Fig. 4.6 on page 103. Finally, the amplitude of the velocity oscillation becomes smaller, since it is suppressed by a factor c_s with respect to the density and c_s is smaller in the presence of baryons. This leads to a suppression of the Doppler contribution to the acoustic peak structure. From Eq. (3.110) we obtain that at the moment of decoupling, $z_{\rm dec} \approx 1100$, we have $R \approx 0.6$.

The solution to (3.111) for time-dependent R can be found in the WKB approximation (Hu & Sugiyama, 1995a), in which case the qualitative picture sketched above slightly changes: the sound speed becomes $k\int c_s {\rm d}\eta$, while the amplitude of the oscillations grows in time as $c_s^{1/2}$. This can be seen simply by considering the quantity $m\omega A^2$, which for an harmonic oscillator is an adiabatic invariant: since in our case the effective mass $m = (1+R)^{1/2}$ decreases in time, it follows that the amplitude $A \sim (1+R)^{-1/4} \sim c_s^{1/2}$.

3.5 Damping

In the above discussion, we have neglected the fact that recombination takes a finite time to complete, and the acoustic oscillations are not frozen instantly. This "finite thickness" of the last scattering surface has a twofold effect: photon diffusion and cancellation. Diffusion damping arises because of the imperfect coupling between photons and baryons, so that photons diffuse out of over-dense into under-dense regions and erase fine scale anisotropies;

cancellation occurs for scales which have the time to oscillate through recombination, so that the effect of photons that last scattered on a crest of the oscillation is cancelled by the contribution of the photons coming from a trough. Cancellation produces a power law damping of the fluctuations (Hu & Sugiyama, 1995a), while diffusion damping is exponential and is by far the dominant effect, and the one to which we now turn our attention. It is often referred to as "Silk damping" (Silk, 1968).

In view of obtaining a dispersion relation $\omega(k)$ for photons accurate to first order in $\dot{\tau}^{-1}$, we look for solutions of the form $V_\gamma \propto \exp \imath \int \omega \mathrm{d}\eta$. At this order we need to include the photon anisotropic stress, which to first order in $\dot{\tau}^{-1}$ from Eq. (2.127c, page 53) is given by (neglecting polarization effects)

$$\Pi_\gamma = \dot{\tau}^{-1} \frac{16}{9} V_\gamma \,. \tag{3.115}$$

Using the anisotropic stress equation (2.51, page 44) we can substitute in the dipole equation (2.127b, page 53) $\Phi = \Psi + \mathcal{H}^2 \Pi_\gamma$. However, we assume that the oscillation time scale is much shorter that the expansion time scale, i.e. $\omega^{-1} \ll \mathcal{H}^{-1}$, so that we can neglect the term $\mathcal{H}^2 \Pi_\gamma$ in the photon dipole. By the same token, in the following we also neglect all time dependencies of the potentials and of R compared with the oscillation time scale.

We now expand the baryon momentum conservation equation (3.109) up to *second order* in $\dot{\tau}^{-1}$, and find, under the above assumptions

$$V_b = V_\gamma - \dot{\tau}^{-1} R(\imath\omega V_\gamma - \Phi) - \dot{\tau}^{-2}(R\omega)^2 V_\gamma + \mathcal{O}(\dot{\tau}^{-3}) \,. \tag{3.116}$$

Inserting this into Eq. (2.127b, page 53) we obtain

$$\imath\omega(1+R)V_\gamma = \frac{1}{4} D_{g,\gamma} + (2+R)\Phi - \dot{\tau}^{-1} V_\gamma \left[(R\omega)^2 - \frac{8}{27} k^2 \right] . \tag{3.117}$$

To lowest order in $\dot{\tau}^{-1}$ we have found in § 3.4 that the quantity $\frac{1}{4} D_{g,\gamma} + (2+R)\Phi$ oscillates with the same frequency as V_γ, see Eq. (3.112). Therefore we set $\frac{1}{4} D_{g,\gamma} + (2+R)\Phi \propto \exp \imath \int \omega \mathrm{d}\eta$, and using the photon monopole equation (3.105) we arrive at

$$\omega^2 = \frac{k^2}{3(1+R)} + \imath \dot{\tau}^{-1} \frac{\omega}{1+R} \left[R^2 \omega^2 + \frac{8}{27} k^2 \right] . \tag{3.118}$$

To zeroth order we find as before $\omega^2 = k^2/[3(1+R)]$, which we can use to obtain the first order solution

$$\omega = \frac{k}{\sqrt{3(1+R)}} + \imath \dot{\tau}^{-1} \frac{k^2}{6(1+R)} \left[\frac{R^2}{(1+R)} + \frac{8}{9} \right] . \tag{3.119}$$

The imaginary term in the frequency induces an exponential damping of the oscillatory solutions of the form $\exp(-k^2/k_\mathrm{D}^2)$, with the characteristic damping scale given by

$$k_\mathrm{D}^{-1} = \int \frac{1}{6\dot{\tau}} \left[\frac{R^2}{(1+R)^2} + \frac{8}{9(1+R)} \right] \mathrm{d}\eta \,. \tag{3.120}$$

Including polarization effects via Eqs. (2.132, page 54) and (2.136, page 54) would increase the damping, by changing the numerical factor 8/9 in the above equation to 16/15.

3.6 Observable quantities

3.6.1 Temperature fluctuations

We now calculate the fluctuations in the CMB photon temperature on the sky. When the photon mean free path becomes larger than the horizon scale, $1/\dot{\tau} \gg 1/\mathcal{H}$, the Universe becomes transparent and photons propagate along null geodesics (*free streaming regime*).

In this section we calculate the photon temperature today with the line of sight method: we formally integrate the Boltzmann equation along the photon path, and obtain the temperature measured today as an integral over a time dependent source term. This approach includes in principle all the effects due to imperfect photons-electrons coupling and reionization as well, and it is the core of the fast numerical algorithms for the integration of the photon Boltzmann equation, such as CMBfast (Seljak & Zaldarriaga, 1996). Another derivation of the same result based on a more physical understanding of the free streaming regime can be found in Durrer (1990).

Consider the collisional Boltzmann equation for the photons temperature $\Theta(\eta, k, \mu = \hat{\mathbf{k}}\cdot\mathbf{n})$ (were we neglect polarization)

$$\dot{\Theta} + \imath k\mu\Theta + \imath k\mu(\Psi + \Phi) = -\dot{\tau}\left[\Theta + \imath\mu k V_b - \Theta_0 - \frac{1}{2}P_2\Theta_2\right], \tag{3.121}$$

and denote with

$$\tau(\eta) \equiv \int_\eta^{\eta_0} \dot{\tau}\mathrm{d}\tilde{\eta} \tag{3.122}$$

the total opacity from the time η until today. Using the equality

$$\frac{\mathrm{d}}{\mathrm{d}\eta}\left(\Theta e^{\imath k\mu\eta}e^{-\tau}\right) = e^{\imath k\mu\eta}e^{-\tau}\left[\dot{\Theta} + \imath k\mu\Theta + \dot{\tau}\Theta\right] \tag{3.123}$$

we obtain from (3.121)

$$\Theta = -\int_0^{\eta_0} e^{\imath k\mu(\eta-\eta_0)}e^{-\tau}\left[\dot{\tau}\left(\imath\mu k V_b - \Theta_0 - \frac{1}{2}P_2\Theta_2\right) + \imath k\mu(\Psi+\Phi)\right]. \tag{3.124}$$

The second term on the right hand side can be integrated by parts and we drop the boundary term, which contributes only to the monopole and is thus unobservable, obtaining

$$\begin{aligned}\Theta(\eta_0, k, \mu) = \int_0^{\eta_0} \mathrm{d}\eta e^{\imath k\mu(\eta-\eta_0)}g(\eta)\left[-\imath\mu k V_b + \Theta_0 + \tfrac{1}{2}P_2\Theta_2 + \Psi + \Phi\right] \\ + \int_0^{\eta_0} \mathrm{d}\eta e^{\imath k\mu(\eta-\eta_0)}e^{-\tau}(\dot{\Psi}+\dot{\Phi}),\end{aligned} \tag{3.125}$$

and we have defined the *visibility function*

$$g(\eta) \equiv \dot{\tau}e^{-\tau}. \tag{3.126}$$

Equation (3.125) is an integral system of equations, since moments $\ell < 3$ of the photons temperature appear on both sides. However, the left hand side can be determined given the time evolution of an handful of quantities which act as a source on the right hand side: the photons moments $\ell < 3$ are calculated from the Boltzmann hierarchy (2.126, page 53), the baryon and CDM velocity and density perturbation from the fluid conservation equations

(2.62–2.63, page 45), while the Bardeen potentials follow from the Poisson equation (2.49) and either the constraint equation (2.50) or the anisotropic stress equation (2.51, page 44). Neutrinos can be included via a collisionless Boltzmann hierarchy, Eq. (2.90, page 48). The great advantage is that only the first few moments of the collisional Boltzmann hierarchy for photons need to be computed accurately in order to obtain the sources of (3.125), reducing the number of coupled differential equations which needs solving from several thousands to a few dozens. This line of sight integration approach is the core algorithm of all modern codes for the computation of the CMB power spectrum (Seljak & Zaldarriaga, 1996).

The visibility function $g(\eta)\mathrm{d}\eta$ in (3.125) encodes the information regarding the ionization history of the Universe, and can be interpreted as the probability that a given CMB photon was last scattered between η and $\eta + \mathrm{d}\eta$. The sharp drop of the free electron density n_e at decoupling makes the visibility function sharply peaked around η_{dec}, cf. the solid line in Fig. 6.15. When the Universe is reionized at later time, the visibility function becomes non-zero again, and the free streaming regime goes once again over in a collisional regime (§ 4.1.3.2).

In the limit of instantaneous recombination, the LSS becomes infinitely thin and the visibility function a delta function peaked at η_{dec}, while we can approximate $e^{-\tau}$ with the Heaviside step function $u(\eta - \eta_{\mathrm{dec}})$. In this limit, the tight coupled fluid approximation for photons goes over directly to the free streaming regime, and there is no generation of photons anisotropic stress nor polarization. Performing the time integral of (3.125) and setting to zeroth order $V_b = V_\gamma$ we find

$$\Theta(\eta_0, k, \mu) = e^{\imath k\mu(\eta_{\mathrm{dec}} - \eta_0)} \left[\Theta^{(\mathrm{OSW})} + \Theta^{(\mathrm{Dpl})} + \Theta^{(\mathrm{ISW})}\right], \tag{3.127}$$

where

$$\begin{aligned} \Theta^{(\mathrm{OSW})} &\equiv [\Theta_0 + \Psi + \Phi](\eta_{\mathrm{dec}}, k) = \left[\frac{1}{4}D_{g,\gamma} + \Phi + \Psi\right](\eta_{\mathrm{dec}}, k) \\ &= \left[\frac{1}{4}D_{s,\gamma} + \Psi\right](\eta_{\mathrm{dec}}, k) \end{aligned} \tag{3.128}$$

$$\Theta^{(\mathrm{Dpl})} \equiv -\imath\mu k V_\gamma(\eta_{\mathrm{dec}}, k) \tag{3.129}$$

$$\Theta^{(\mathrm{ISW})} \equiv \int_{\eta_{\mathrm{dec}}}^{\eta_0} \mathrm{d}\eta e^{\imath k\mu(\eta - \eta_0)}(\dot{\Psi} + \dot{\Phi})(\eta, k) \tag{3.130}$$

The temperature fluctuation consists of three terms:

- **The ordinary Sachs-Wolfe (OSW)** part, $\Theta^{(\mathrm{OSW})}$. The photons temperature monopole Θ_0 on the last scattering surface, together with the potential terms Φ and Ψ, reflect intrinsic inhomogeneities in the radiation fluid and in the metric at the moment of decoupling. On large scales, the ordinary SW effect is responsible for the SW plateau in the temperature power spectrum, while on intermediate scales the oscillations of $D_{g,\gamma}$ produce the familiar peak structure.

- **The Doppler term** $\Theta^{(\mathrm{Dpl})} \propto kV_b$ arises because of the relative velocity of observer and emitter. Its contribution shows up on the acoustic peak scale.

- **The integrated Sachs-Wolfe (ISW)** effect produces the term $\Theta^{(\mathrm{ISW})}$, and it is induced by a time dependence of the Bardeen potentials along the path of the photons.

The *early* ISW effect is due to the fact the the universe is not completely matter dominated at recombination and therefore the potentials are not exactly constant; the *late* ISW is generated when the late universe becomes dominated by the curvature or a cosmological constant term, both of which induce a time dependence in the potentials.

The dependence of the anisotropies on the cosmological parameters is presented in § 4.1.

3.6.2 Angular power spectra

The relevant quantities for the comparison of theoretical models and observations are the temperature and polarization angular power spectra, which we introduce in this section. We refer the reader to § 5.1.1 for precise definitions of the terminology. We denote by $\langle\cdot\rangle$ the theoretical ensemble average over realizations.

Temperature power spectrum

The temperature fluctuation in direction $\mathbf{n}$ on the sky measured by an observer today (η_0) and here ($\mathbf{x_0}$) is a superposition of plane wave contributions (in a flat Universe)

$$\Theta(\eta_0, \mathbf{x_0}, \mathbf{n}) = \frac{1}{(2\pi)^{3/2}} \int \mathrm{d}^3k \; \Theta(\eta_0, \mathbf{k}, \mathbf{n}) e^{\imath \mathbf{x_0 k}} \tag{3.131}$$

and each Fourier mode can be expanded in spherical harmonics on the 2-sphere as

$$\Theta(\eta_0, \mathbf{k}, \mathbf{n}) = \sum_{\ell=0}^{\infty} \sum_{m=-\ell}^{\ell} a_{\ell m}(\mathbf{k}, \eta_0) Y_{\ell m}(\mathbf{n}) \,, \tag{3.132}$$

where the expansion coefficients $a_{\ell m}(\mathbf{k})$ are given by

$$a_{\ell m}(\mathbf{k}) = \int \mathrm{d}\Omega_{\mathbf{n}} \Theta(\mathbf{k}, \mathbf{n}) Y_{\ell m}(\mathbf{n}) \tag{3.133}$$

$$= 4\pi \Theta_\ell(\mathbf{k}) Y_{\ell m}(\hat{\mathbf{k}}) \,. \tag{3.134}$$

In deriving the last expression we have expanded the temperature fluctuation in Legendre polynomials as in (2.84, page 48) and used the addition theorem and orthogonality relation for spherical harmonics:

$$\sum_{m=-\ell}^{\ell} Y_{\ell m}(\mathbf{n}) Y^*_{\ell m}(\mathbf{n}') = \frac{2\ell+1}{4\pi} P_\ell(\mathbf{n} \cdot \mathbf{n}') \,, \tag{3.135}$$

$$\int \mathrm{d}\Omega_{\mathbf{n}} Y_{\ell m}(\mathbf{n}) Y^*_{\ell' m'}(\mathbf{n}) = \delta_{\ell \ell'} \delta_{m m'} \,. \tag{3.136}$$

We can perform the harmonic expansion (3.132) directly in real space rather than in Fourier space, with coefficients $a_{\ell m}(\mathbf{x_0})$ (for which we will neglect the argument $\mathbf{x_0}$ from now on), obviously related to $a_{\ell m}(\mathbf{k})$ by

$$a_{\ell m} = \frac{1}{(2\pi)^{3/2}} \int \mathrm{d}^3 \, \mathbf{k} a_{\ell m}(\mathbf{k}) e^{\imath \mathbf{k x_0}} \,. \tag{3.137}$$

We are interested in the 2-point temperature correlation function C on the sky between two directions $\mathbf{n}$ and $\mathbf{n}'$. By choosing our coordinate system in such a way that the direction

$\mathbf{n}$ corresponds to the z-axis, and introducing spherical coordinates we can write $\mathbf{n}' = (\phi, \vartheta)$ and $\mathbf{n} \cdot \mathbf{n}' = \cos(\vartheta)$. If we assume statistical homogeneity and isotropy for the random field Θ, see § 5.1.1, then the correlation function does not depend on the observer's position (homogeneity) nor on the azimutal angle ϕ (isotropy). Therefore

$$\begin{aligned} C(\vartheta) &\equiv \langle \Theta(\eta_0, \mathbf{x_0}, \mathbf{n}) \cdot \Theta(\eta_0, \mathbf{x_0}, \mathbf{n}') \rangle \\ &= \frac{1}{4\pi} \sum_\ell (2\ell + 1) C_\ell P_\ell(\mathbf{n} \cdot \mathbf{n}') \,, \end{aligned} \tag{3.138}$$

where we have defined the *CMB angular power spectrum* by

$$\langle a_{\ell m} \cdot a^*_{\ell' m'} \rangle = \delta_{\ell\ell'} \delta_{mm'} C_\ell \,. \tag{3.139}$$

The fact that C_ℓ does not depend on $\mathbf{x_0}$ is a consequence of the assumption of homogeneity, while isotropy requires that it does not depend on the index m, which would introduce an azimutal dependence. It is also customary to assume that the $a_{\ell m}$'s are Gaussian random fields, as motivated by inflation, but this is not strictly necessary at this stage. Eq. (3.138) shows that the angular power spectrum is the harmonic transform of the correlation function on the 2-sphere and for Gaussian variables it contains the full statistical information. If the $a_{\ell m}$'s are Gaussian distributed, then the Fourier coefficients $a_{\ell m}(\mathbf{k})$ are Gaussian random variables as well. From the assumption of homogeneity it follows that $\langle a_{\ell m}(\mathbf{k}) \rangle = \delta^{(\mathrm{D})}(\mathbf{k})$, where $\delta^{(\mathrm{D})}$ denotes the Dirac delta function. Homogeneity and isotropy together imply that

$$\langle |a_{\ell m}|^2 \rangle = \frac{1}{(2\pi)^3} \int \mathrm{d}^3\mathbf{k} \langle |a_{\ell m}(\mathbf{k})|^2 \rangle \,. \tag{3.140}$$

We now relate the angular power spectrum to the temperature multipoles: this is done by observing that the evolution equations (2.126, page 53) for Θ_ℓ are independent of $\hat{\mathbf{k}}$, and therefore we can write

$$\Theta_\ell(\eta, \mathbf{k}) = \Theta_\ell(\eta, k) \chi(\mathbf{k}) \,, \tag{3.141}$$

where we assume that $\chi(\mathbf{k})$ are the Fourier components of a Gaussian, isotropic and homogeneous random field. As a consequence

$$\langle \chi(\mathbf{k}) \cdot \chi^*(\mathbf{k}') \rangle = \delta^{(\mathrm{D})}(\mathbf{k} - \mathbf{k}') \, \langle |\chi(k)|^2 \rangle \,. \tag{3.142}$$

Now from (3.139) and using Eqs. (3.141), (3.140) and (3.134) we obtain

$$C_\ell = 4\pi \int \frac{\mathrm{d}k}{k} \, P_\chi(k) \, |\Theta_\ell(\eta_0, k)|^2 \,. \tag{3.143}$$

We shall later identify χ with the primordial curvature or entropy perturbation, see Eq. (4.5, page 91), and call

$$P_\chi(k) \equiv \frac{k^3}{2\pi^2} \langle |\chi|^2 \rangle \tag{3.144}$$

the curvature (or entropy) power spectrum: this quantity gives the contribution to C_ℓ per logarithmic k-interval of the primordial fluctuation.

The *photons transfer function* $\Theta_\ell(\eta_0, k)$ in Eq. (3.143) above is an intrinsically 2-dimensional quantity which gives information about how the initial power is mapped onto the angular

power spectrum. It can be evaluated from Eq. (3.125, page 79), by observing that the angle $\mu = \hat{\mathbf{k}} \cdot \mathbf{n}$ in the integrand can be eliminated by replacing

$$e^{\imath k\mu(\eta-\eta_0)} g \imath k \mu V_b = \frac{\mathrm{d}}{\mathrm{d}\eta}\left(e^{\imath k\mu(\eta-\eta_0)} g V_b\right) - e^{\imath k\mu(\eta-\eta_0)} \dot{g} V_b - e^{\imath k\mu(\eta-\eta_0)} g \dot{V}_b \tag{3.145}$$

and dropping the total derivative which only gives an unobservable monopole term. Therefore we can rewrite (3.125, page 79) as

$$\Theta(\eta_0, k, \mu) = \int_0^{\eta_0} \mathrm{d}\eta e^{\imath k\mu(\eta-\eta_0)} \mathcal{S}(\eta, k) \tag{3.146}$$

with the source term of the form

$$\begin{aligned} \mathcal{S}(\eta, k) = g\left[\frac{\dot{V}_b}{k} + \Theta_0 - \frac{\Theta_2}{4} + \Psi + \Phi\right] - \dot{g}\left[\frac{V_b}{k} + \frac{3}{4}\frac{\Theta_2}{k^2}\right] - \ddot{g}\frac{3}{4}\frac{\dot{\Theta}_2}{k^2} \\ + e^{-\tau}(\dot{\Psi} + \dot{\Phi})\,. \end{aligned} \tag{3.147}$$

Now we expand the plane wave in radial and angular eigenfunctions, Bessel functions and Legendre polynomials respectively, using the Rayleigh formula

$$e^{\imath k\mu(\eta-\eta_0)} = \sum_\ell \imath^\ell (2\ell+1) j_\ell(k(\eta_0 - \eta)) P_\ell(\mu)\,, \tag{3.148}$$

and we obtain for the temperature transfer function

$$\Theta_\ell(\eta_0, k) = \imath^\ell \int_0^{\eta_0} \mathrm{d}\eta \mathcal{S}(\eta, k) j_\ell(k(\eta_0 - \eta))\,. \tag{3.149}$$

This is shown in the top panels of Fig. 3.1 for adiabatic and isocurvature CDM initial conditions.

Together, Eqs. (3.149) and (3.143) allow the computation of the CMB angular power spectrum and neatly split the geometric effects from the physics: all of the dynamical evolution is encoded in the source function $\mathcal{S}(\eta, k)$, while the Bessel function accounts for the projection from 3-dimensional k-space on the 2-sphere. The generalization of this result for the $\mathcal{K} \neq 0$ case can be found in Zaldarriaga et al. (1998); Zaldarriaga & Seljak (2000); Lewis et al. (2000). The temperature and E-polarization spectra of a concordance model for adiabatic and isocurvature CDM initial conditions are displayed in the top left panel of Fig. 4.9 on page 106.

Polarization power spectrum

As mentioned in § 2.2.5.5, polarization of scalar modes is conveniently described by the E polarization mode, supplemented by the cross-correlator between E and T (temperature). As for temperature, we can formally integrate the Boltzmann equation for the Stokes parameter Q, Eq. (2.131, page 54), along the line of sight and obtain

$$\Theta^Q(\eta_0, k, \mu) = -\frac{1}{2}\int_0^{\eta_0} e^{\imath k\mu(\eta-\eta_0)} g(\eta)\,(1 - P_2)\left(\Theta_2 + \Theta_2^Q - \Theta_0^Q\right)\,. \tag{3.150}$$

The E-polarization power spectrum and the ET-correlator (superscript C) are defined as

$$\langle a_{\ell m}^E \cdot a_{\ell' m'}^{*E} \rangle = \delta_{\ell\ell'}\delta_{mm'} C_\ell^E\,, \tag{3.151}$$

$$\langle a_{\ell m}^T \cdot a_{\ell' m'}^{*E} \rangle = \delta_{\ell\ell'}\delta_{mm'} C_\ell^C\,, \tag{3.152}$$

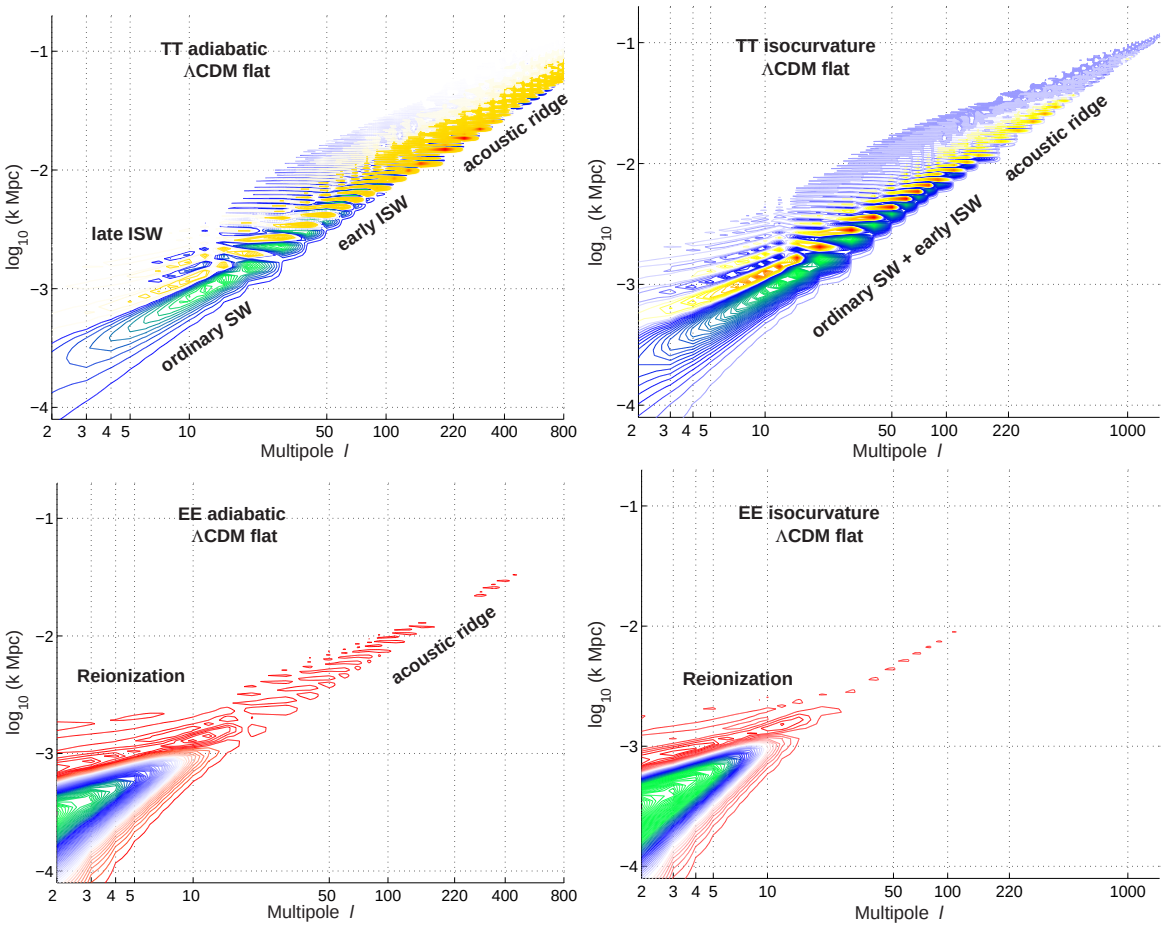

Figure 3.1: Temperature (top) and polarization (bottom) transfer function $\Theta_\ell(\eta_0, k)$ and $\Delta^E_\ell(\eta_0, k)$ for adiabatic (left panels) and isocurvature CDM (right panels) initial conditions. The color scales are arbitrary, and have been chose as too highlight the features of the transfer functions. In particular, the color coding is not in scale between the different plots.

and in analogy with the treatment for the temperature spectrum they can be computed as a superposition of k modes of a source function integrated over time:

$$C^E_\ell = 4\pi \int \frac{\mathrm{d}k}{k} P_\chi(k) \, |\Delta^E_\ell(\eta_0, k)|^2 \,, \tag{3.153}$$

$$\Delta^E_\ell(\eta_0, k) = \sqrt{\frac{(\ell+2)!}{(\ell-2)!}} \int_0^{\eta_0} \mathrm{d}\eta \mathcal{S}^E(\eta, k) j_\ell(k(\eta_0 - \eta)) \,, \tag{3.154}$$

$$\mathcal{S}^E(\eta, k) = \frac{3g(\eta)}{4k^2(\eta_0 - \eta)^2} \left(\Theta_2 + \Theta^Q_2 - \Theta^Q_0\right) \,. \tag{3.155}$$

The cross-correlator spectrum is computed using (3.149) as

$$C^C_\ell = 4\pi \int \frac{\mathrm{d}k}{k} P_\chi(k) \, \Theta^*_\ell(\eta_0, k) \Delta^E_\ell(\eta_0, k) \,. \tag{3.156}$$

The polarization transfer function $\Delta^E_\ell(\eta_0, k)$ is plotted in Fig. 3.1 for adiabatic and isocurvature CDM initial conditions.

The degree of polarization is proportional to the magnitude of the temperature quadrupole at last scattering. Since during the tight coupling regime the temperature quadrupole cannot grow, polarization is generated in the relatively short transition between the strong coupling and the free streaming regime. To first order in $\dot{\tau}^{-1}$, the temperature quadrupole is proportional to the temperature dipole, see (4.30, page 95). The polarization amplitude is thus proportional to the temperature dipole at recombination times the width of the last scattering surface (Zaldarriaga & Harari, 1995), resulting in a polarization signal two orders of magnitude lower than the temperature signal.

3.6.3 Matter power spectrum

Let $\delta(\eta, \mathbf{x})$ denote the real-space density contrast in the matter component in the comoving gauge; hence δ corresponds to the gauge invariant variable Δ_m defined in Eq. (2.37, page 42). We will drop the time dependence when not needed, and write δ instead of Δ_m to simplify the notation. For clarity, the Fourier transform of the variables is denoted by a subscript "$\mathbf{k}$", in this section only.

The real space correlation function is defined as

$$\xi(\mathbf{r}) \equiv \langle \delta(\mathbf{x}) \cdot \delta(\mathbf{x}+\mathbf{r}) \rangle \,, \tag{3.157}$$

where $\langle \cdot \rangle$ denotes an average over realizations, see § 5.1.1 for precise definitions. It is the expectation value of $\delta_2 = \delta(\mathbf{x}+\mathbf{r})$ and $\delta_1 = \delta(\mathbf{x})$ under the 2-point probability distribution function for δ_1, δ_2. We write $\delta(\mathbf{x})$ as

$$\delta(\mathbf{x}) = \frac{1}{(2\pi)^{3/2}} \int \mathrm{d}^3\mathbf{k} \delta_{\mathbf{k}} e^{\imath \mathbf{k}\mathbf{x}} \tag{3.158}$$

where we denote by $\delta_{\mathbf{k}}$ the Fourier transform (in flat space) of $\delta(\mathbf{x})$. We postulate that $\delta(\mathbf{x})$ is a Gaussian distributed, isotropic and homogeneous random field, see § 5.1.1, and therefore the quantity $\langle \delta^*_{\mathbf{k}} \cdot \delta_{\mathbf{k}'} \rangle$ vanishes for $\mathbf{k} \neq \mathbf{k}'$ (homogeneity) and it only depends on the modulus, not the direction of $\mathbf{k}$ (isotropy):

$$\langle \delta^*_{\mathbf{k}} \cdot \delta_{\mathbf{k}'} \rangle = (2\pi)^{3/2} \delta^{(\mathrm{D})}(\mathbf{k} - \mathbf{k}') \; P_m(k) \tag{3.159}$$

where $\delta^{(\mathrm{D})}$ denotes the Dirac delta function. We call $P_m(k)$ the *matter power spectrum*. Replacing (3.158) in (3.157) we obtain

$$\xi(r) = \frac{1}{(2\pi)^{3/2}} \int \mathrm{d}^3\mathbf{k} P_m(k) e^{\imath \mathbf{k}\mathbf{r}} = \frac{2}{\sqrt{2\pi}} \int \mathrm{d}k k^2 \frac{\sin rk}{rk} P_m(k) \,, \tag{3.160}$$

showing that the correlation function is the Fourier transform of the matter power spectrum.

Our aim is to compute the power spectrum today as a function of the spectral distribution in the early Universe in the adiabatic CDM scenario. To this end, we make use of the results of linear perturbation theory presented in the previous sections for the growth of matter perturbations in a Universe containing CDM and photons only. Clearly, these computations are valid only as long as the scale considered is in the linear regime, i.e. $\delta_{\mathbf{k}} \ll 1$. We only sketch the elements which are needed in the following, referring the reader to e.g. Peebles (1980); Padmanabhan (1993); Liddle & Lyth (2000) for a full account.

Perturbations $\delta_{\mathbf{k}}$ over a comoving length $\lambda \sim k^{-1}$ behave differently depending whether they are outside ($k < \mathcal{H}$) or inside ($k > \mathcal{H}$) the Hubble length. For a given scale k, we denote by $\eta_{\rm ent}$ the time at which that scale crosses inside the horizon, i.e. $\mathcal{H}(\eta_{\rm ent}) = k$ and by $k_{\rm eq}$ the wavelength which enters the horizon at the time of matter-radiation equality, i.e. $k_{\rm eq} = \mathcal{H}(\eta_{\rm eq})$. We thus need to distinguish two cases: scales $k > k_{\rm eq}$ enter the horizon in the radiation dominated epoch, while $k < k_{\rm eq}$ enter the horizon after matter domination. We shall restrict ourselves to length scales λ which are large enough not to be wiped out by free streaming, i.e. $\lambda > \lambda_{\rm FS}$, see Padmanabhan (1993) for details.

For $k > k_{\rm eq}$ and $\eta_{\rm ent} < \eta < \eta_{\rm eq}$, $\delta_{\mathbf{k}}(\eta)$ stays approximately constant after horizon crossing because the radiation dominated epoch suppresses the growth of perturbation in a dust-like component; this is called the Meszaros effect (Meszaros, 1974). For $\eta > \eta_{\rm eq}$ the Universe is matter dominated and the situation is analogous to the single fluid case examined in § 3.1, and the perturbation grows as $\delta_{\mathbf{k}} \propto a$, see Eq. (3.14, page 61). Wavelengths which enter the horizon in the matter dominated epoch, $k < k_{\rm eq}$, start growing as soon as they cross the horizon, $\delta_{\mathbf{k}}(\eta) \propto a$ for $\eta > \eta_{\rm ent}$, by the same argument given above. Summarizing, we have that

$$\delta_{\mathbf{k}}(\eta > \eta_{\rm ent}) \propto \begin{cases} \delta_{\mathbf{k}}(\eta_{\rm ent}) \dfrac{a}{a_{\rm eq}} & \text{for } k > k_{\rm eq} \\ \delta_{\mathbf{k}}(\eta_{\rm ent}) \dfrac{a_{\rm eq}}{a_{\rm ent}} \dfrac{a}{a_{\rm eq}} & \text{for } k < k_{\rm eq} \end{cases} , \tag{3.161}$$

and therefore we know $\delta_{\mathbf{k}}$ for all subsequent times once we specify $\delta_{\mathbf{k}}(\eta_{\rm ent})$, the value of the density contrast for the wavelength k *at the moment when that wavelength crossed inside the horizon.* Since for a given wavelength $\eta_{\rm ent} \propto 1/k$, horizon crossing happens at a different time for each scale. We notice that in the second line of Eq. (3.161) we can rewrite the factor $a_{\rm eq}/a_{\rm ent}$ as

$$\frac{a_{\rm eq}}{a_{\rm ent}} = \left(\frac{\eta_{\rm eq}}{\eta_{\rm ent}}\right)^2 = \left(\frac{k}{k_{\rm eq}}\right)^2 \propto k^2 , \tag{3.162}$$

where in the first equality we have used the fact that $a \propto \eta^2$ in the matter dominated universe.

Given that the range of scales of cosmological interest is not too wide, we can make the following power law Ansatz for the scale dependence of the perturbation at horizon crossing

$$\delta_{\mathbf{k}}(\eta_{\rm ent}) = A k^{-\alpha} . \tag{3.163}$$

An important quantity is $k^3/(2\pi)^{3/2} P_m(k)$, which from (3.160) gives the contribution per logarithmic k-interval to the real space correlation function, and which with the above Ansatz evaluates to

$$\left.\frac{k^3}{(2\pi)^{3/2}} P_m(k)\right|_{\eta_{\rm ent}} \propto k^{3-2\alpha} = \text{ const for } \alpha = 3/2 . \tag{3.164}$$

This quantity can also be interpreted as the variance of the mass contained in spheres of diameter $\lambda \sim 1/k$ at horizon crossing, see e.g. Padmanabhan (1993); for the value $\alpha = 3/2$ the variance is the same on all scales.

We might prefer to specify our Ansatz not at horizon crossing, but rather for some fixed initial time (the same for all scales) η_i. In order to relate $\delta_{\mathbf{k}}(\eta_i)$ with $\delta_{\mathbf{k}}(\eta_{\rm ent})$, we notice that on super-horizon scales $k < \mathcal{H}$ and for times $\eta_{\rm ent} > \eta > \eta_{\rm eq}$ we have $\delta_{\mathbf{k}} \propto a \propto \eta^2$ from

Eq. (3.14, page 61). For the case $k < \mathcal{H}$ in the radiation epoch, $\eta < \eta_{\rm ent} < \eta_{\rm eq}$ we can use the adiabatic solution (3.26, page 63) and the relation

$$\delta_{\mathbf{k}} \equiv \Delta_m = \frac{3+3a}{4+3a} D + \frac{4}{4+3a} S \approx \frac{3}{4} D + S \propto a^2 \propto \eta^2 \,, \tag{3.165}$$

and the approximation is valid for $a < a_{\rm eq}$. In conclusion, the comoving dark matter density contrast grows as η^2 at all epochs while outside the horizon. Therefore we obtain (with $\eta_{\rm ent} > \eta_i$ for all scales of interest)

$$\delta_{\mathbf{k}}(\eta_{\rm ent}) = \left(\frac{\eta_{\rm ent}}{\eta_i} \right)^2 \delta_{\mathbf{k}}(\eta_i) \propto k^{-2} \delta_{\mathbf{k}}(\eta_i) \,. \tag{3.166}$$

It is customary to make a power law Ansatz for the matter power spectrum at the time η_i of the form

$$P_m(k, \eta_i) = B k^n \tag{3.167}$$

and by the relation (3.166) the index n is related to α by

$$n = -2\alpha + 4 \,. \tag{3.168}$$

The value $\alpha = 3/2$ which yields a constant-mass-variance on all scales at horizon crossing corresponds to $n = 1$, the so-called "scale invariant spectral index", also known as Harrison-Zel'dovich spectrum (Harrison, 1970; Zel'dovich, 1972). The power spectrum today then becomes in terms of n, from (3.161)

$$\delta_{\mathbf{k}}(\eta_0) \propto \begin{cases} k^{n-4} & \text{for } k > k_{\rm eq} \\ k^n & \text{for } k < k_{\rm eq} \end{cases} \,. \tag{3.169}$$

The length scale which crosses the horizon at equality, $\lambda_{\rm eq} \approx 13/(\Omega_m h^2)$ Mpc corresponds to a peak in the power spectrum: fluctuations on larger scales, $k < k_{\rm eq} \sim 1/\lambda_{\rm eq}$ retain their primordial shape, while perturbations on smaller scales have their spectrum multiplied by k^{-4}. The above arguments only apply in the linear region, i.e. for $k \lesssim 0.3\ h/\text{Mpc}$, above which non-linear growth of the fluctuations invalidate perturbation theory and a full numerical simulation is required to follow the evolution.

Finally, we can easily relate the matter power spectrum to the Bardeen potential by using the Poisson equation (2.49, page 44). If we consider the value of $\Psi_{\mathbf{k}}(\eta_{\rm ent})$, the Fourier transform of Ψ evaluated at horizon crossing, we have from the Poisson equation, noticing that $\mathcal{H}(\eta_{\rm ent}) = 1/k$, $\delta_{\mathbf{k}} = \Delta_m \sim \Delta_\gamma \sim D$ by the adiabaticity condition, that $\Psi_{\mathbf{k}}(\eta_{\rm ent}) \sim -\delta_{\mathbf{k}}(\eta_{\rm ent})$. Therefore for the power spectrum of the Bardeen potential, defined as

$$P_\Psi \equiv \frac{k^3}{2\pi^2} \langle |\Psi_{\mathbf{k}}|^2 \rangle \tag{3.170}$$

we have that

$$P_\Psi(k)\big|_{\eta_{\rm ent}} \propto k^3 P_m(k)\big|_{\eta_{\rm ent}} \propto k^{n-1} \,, \tag{3.171}$$

and the $n = 1$ scale invariant spectrum corresponds to $P_\Psi(\eta_{\rm ent}) =$ const. Or we can specify P_Ψ at a fixed initial time η_i, in which case we obtain again from the Poisson equation

$$P_\Psi(k)\big|_{\eta_i} \propto k^{-1} P_m(k)\big|_{\eta_i} \propto k^{n-1} \,. \tag{3.172}$$

The fact that there is no evolution in the power spectrum of Ψ until horizon crossing is of course a consequence of the fact that $\Psi_{\mathbf{k}} \approx$ const on super-horizon scales, as shown in § 3.2. The same scaling applies for the power spectrum of the gauge invariant curvature perturbation ζ, which is constant on super-horizon scales for adiabatic perturbations, and proportional to Ψ.

Chapter 4

Parameter dependence

This chapter presents a brief review of the dependence of the CMB power spectra on the standard cosmological parameters and on general initial conditions, building on the results of the previous sections. Understanding the impact of the parameters on the observable spectra builds the framework for parameter extraction from data, which is the subject of Part III.

In § 4.1 we concisely review the origin and main parameters dependencies of well known features of the power spectrum: the large scale Sachs-Wolfe plateau, the acoustic oscillations, and the damping tail. Introductory reviews on this topic can be found in e.g. Kosowsky (2002) and Hu et al. (1997). A detailed physical understanding in a fully analytical approach is explained in Hu & Sugiyama (1995a,b, 1996). In view of efficient and accurate parameter estimation, fundamental degeneracies in the CMB spectra are best understood by introducing a set of analytical functions of the parameters which the CMB probes directly, and upon which the spectra dependence is almost linear (Kosowsky et al., 2002). We call this new basis in parameter space "normal parameters set", and we illustrate it in § 4.2.

In § 4.3 the CMB angular power spectra for general isocurvature initial conditions in a Universe containing CDM, baryons, photons and neutrinos are presented. The four modes adiabatic, CDM isocurvature, neutrino density and neutrino velocity – along with a baryon isocurvature mode which is equal to the CDM mode up to a rescaling constant – span the whole space of non-diverging solutions of Einstein's equations at early times (Bucher et al., 2000), and thus their superposition constitutes the most general type of initial conditions for CMB anisotropy.

4.1 Standard parameters

The detailed shape of the CMB temperature and polarization spectra depends on the value of the cosmological parameters and on the type of initial conditions in characteristic ways. However, certain combination of parameters lead to very similar spectra: this causes degeneracies among some parameters, which cannot be reconstructed with CMB alone, but require the inclusion of external data-sets.

Polarization information helps breaking temperature degeneracies because of two characteristic features: the first is that after decoupling the polarization state is preserved by free streaming, and the polarization spectrum is only modified by rescattering due to reionization (§ 4.1.3.2). Therefore in a sense polarization is a more clean probe of the decoupling than

temperature. The second reason is that while the acoustic peaks in temperature are dominated by the monopole of the temperature fluctuation on the LSS, the peaks in E-polarization reflect the dipole component at decoupling, i.e. the photon bulk velocity (§ 4.1.2.1).

In the following we revisit the main parameter dependence of the CMB spectra: for the sake of illustrating the physical effects involved, we divide the CMB power spectrum in three distinct regions, corresponding to different angular separations on the sky with the approximate relation $\vartheta \sim \pi/\ell$.

- **Large scales:** on scales larger than the Hubble radius at decoupling, $k\eta_{\rm dec} \ll 1$, perturbations are dominated by the ordinary Sachs-Wolfe effect, given by the combination of the intrinsic temperature fluctuations on the LSS and the gravitational redshift induced by climbing out of the potential well. In non-flat cosmologies, or models with a considerable value of the cosmological constant, the late ISW effect also contributes. This region corresponds roughly to the COBE scale, $\ell \lesssim 30$ and $\vartheta \gtrsim 7^\circ$.

  Reionization produces a a characteristic increase of E-polarization on large scales, the so-called "polarization bump".

- **Acoustic region:** inside the sound horizon photon pressure cannot be neglected, and scales within the sound horizon $k\int c_s\eta \gtrsim 1$ oscillate, while gravitational infall becomes negligible because of potential decay inside the horizon. On intermediate scales $50 \lesssim \ell \lesssim 600$ the CMB power spectrum displays a rich peak structure, reflecting the contributions of density oscillations and Doppler term on the LSS. The early ISW effect contributes at roughly the 20% level up to the first acoustic peak (for adiabatic models). Those scales have a typical angular separation on the sky ranging from about 10° down to a few $10'$.

- **Damping tail:** wavelengths smaller than the diffusion damping scale $1/k_{\rm D}$ given in (3.120, page 78) are exponentially suppressed and this causes a drop in power above $\ell \sim 800$ or $\vartheta \lesssim 1'$. This effect combines with rescattering due to reionization, which also erases fine-scale anisotropies.

4.1.1 Large scales

We wish to investigate the expected temperature fluctuations on very large scales in the general case of a superposition of primordial adiabatic and isocurvature CDM initial conditions. We look at wavelength $k \ll k_{\rm dec}$ which at decoupling where still outside the horizon and we consider a zeroth order approximation which neglects any anisotropic stress and the baryon influence (i.e. set $R = 0$). If we take decoupling to happen well into matter domination, we can also neglect the ISW contribution since the potentials are equal and constant – see Eq. (3.14, page 61) – and to this level of approximation we can set $V_b = V_\gamma$. With this approximations we have for each Fourier mode from Eqs. (3.128, page 80) and (3.129, page 80)

$$\Theta(\eta_0, k, \mu) = e^{\imath k\mu(\eta_{\rm dec}-\eta_0)}\left[\frac{1}{4}D_{g,\gamma} + 2\Phi - \imath k\mu V_\gamma\right](\eta_{\rm dec}, k)\,. \tag{4.1}$$

In the adiabatic case, we can neglect the contribution of the Doppler term which behaves as a sine and hence disappears on large scales, $k\eta_{\rm dec} \ll 1$, while the cosine oscillation of the

density perturbation $D_{g,\gamma}$ becomes constant, see (3.56, page 67). Therefore for adiabatic initial conditions, from the solution (3.56) it follows

$$\Theta(\eta_0, k, \mu) \approx e^{\imath k\mu(\eta_{\rm dec}-\eta_0)} \left[\left(\frac{1}{3}\Phi_{\rm MD} - 2\Phi_{\rm MD}\right) + 2\Phi_{\rm MD}\right] \quad \text{(adiabatic)}, \tag{4.2}$$

where $\Phi_{\rm MD}$ denotes the value of Φ at decoupling well within matter domination. On the right hand side, the term $-2\Phi_{\rm MD}$ comes from the solution (3.56), and its negative sign reflects the fact that the temperature is larger inside potential wells ($\Phi < 0$), so that photons are blushifted when they fall *into* the well. The term $2\Phi_{\rm MD}$ represents the gravitational redshift which photons experience when they climb *out* of the potential as they free stream after decoupling, which exactly cancels the gravitational blueshift term in the absence of baryons. In conclusion we have

$$\Theta(\eta_0, k, \mu) \approx e^{\imath k\mu(\eta_{\rm dec}-\eta_0)} \frac{1}{3}\Phi_{\rm MD} \quad \text{(adiabatic)}. \tag{4.3}$$

For isocurvature initial conditions, we have that $D_{g,\gamma}(\eta_{\rm dec}) = 0$, which follows from (3.48, page 66) with the isocurvature condition $\Phi_0 = 0$. The Doppler term can again be neglected with respect to the potential, because from (3.53, page 67) we have that $kV_\gamma \sim k/\mathcal{H}\Phi \ll \Phi$ and (4.1) reduces to

$$\Theta(\eta_0, k, \mu) \approx e^{\imath k\mu(\eta_{\rm dec}-\eta_0)} 2\Phi_{\rm MD} \quad \text{(isocurvature)}, \tag{4.4}$$

the well-known result that isocurvature initial conditions produce large scale fluctuations six times larger than in the adiabatic case *for the same value of the Bardeen potential* on the last scattering surface.

More interestingly, we can relate the large-scale temperature fluctuations to the amplitude of the primordial curvature and entropy spectra. Rewriting (4.3–4.4) in terms of the curvature and entropy perturbations in the radiation era via Eqs. (3.33–3.36, page 65), yields for the source term (3.147, page 83)

$$\mathcal{S}(\eta, k) = \delta(\eta - \eta_{\rm dec}) \left[\frac{\zeta_0}{5}\psi(k) - \frac{2}{5}S_0\phi(k)\right], \tag{4.5}$$

where $\psi(k)$ and $\phi(k)$ are the Fourier components of random fields which we assume are Gaussian distributed, isotropic and homogeneous, see § 5.1.1, evaluated at some initial time η_i deep in the radiation epoch. For their *power spectrum* we make a power low Ansatz

$$P_\psi(k)\big|_{\eta_i} \equiv \frac{k^3}{2\pi^2}\langle|\psi(k)|^2\rangle = \zeta_0^2 \left(\frac{k}{k_{\rm P}}\right)^{n_{\rm s}-1}, \tag{4.6}$$

$$P_\phi(k)\big|_{\eta_i} \equiv \frac{k^3}{2\pi^2}\langle|\phi(k)|^2\rangle = S_0^2 \left(\frac{k}{k_{\rm P}}\right)^{n_e-1}, \tag{4.7}$$

$$P_c(k)\big|_{\eta_i} \equiv \frac{k^3}{2\pi^2}\langle\psi(k)\cdot\phi^*(k)\rangle = \zeta_0 S_0 \left(\frac{k}{k_{\rm P}}\right)^{n_c-1} \cos(\Delta_c)\,. \tag{4.8}$$

The constants ζ_0 and S_0 are dimensionless and positive, while the angle Δ_c parameterizes the correlation between entropy and isocurvature perturbations; the constant $k_{\rm P}$ is a pivot scale, for which a popular choice is $k_{\rm P} = 0.05$ Mpc^{-1}, and we have defined $n_c \equiv (n_{\rm s} + n_e)/2$.

The power law index $n_{\rm s}$ is the *scalar spectral index*: $n_{\rm s} \approx 1$ is a generic prediction of inflation, almost independently of the particular model, and is called "scale-invariant" or Harrison-Zel'dovich (Harrison, 1970; Zel'dovich, 1972) spectral index. The reason for the name is explained in § 3.6.3. Since $\Psi \propto \zeta$ up to constant factors, Ψ and ζ have the same spectrum.

From (3.143, page 82) the angular power spectrum on large scales ($\ell \lesssim 20$) is then given by

$$C_\ell = 4\pi \int \frac{{\rm d}k}{k} \left[\frac{\zeta_0^2}{25} \left(\frac{k}{k_{\rm P}} \right)^{n_{\rm s}-1} + \frac{4S_0^2}{25} \left(\frac{k}{k_{\rm P}} \right)^{n_e-1} - \frac{4}{25} \zeta_0 S_0 \cos(\Delta_c) \left(\frac{k}{k_{\rm P}} \right)^{n_{c-1}} \right] \times \\ \times j_\ell^2 \left(k(\eta_0 - \eta_{\rm dec}) \right) . \tag{4.9}$$

The integral can be performed analytically provided all the indexes are within the range $-3 < n_X < 3$ and in the approximation $k(\eta_0 - \eta_{\rm dec}) \approx k\eta_0$ (Gradshteyn & Ryzhik, 1965). The result is

$$C_\ell = 2\pi^2 \left[\frac{\zeta_0^2}{25} f(n_{\rm s}, \ell) + \frac{4S_0^2}{25} f(n_e, \ell) - \frac{4}{25} \zeta_0 S_0 \cos(\Delta_c) f(n_c, \ell) \right] . \tag{4.10}$$

The function f contains the dependence on the spectral indexes, and it is given by

$$f(n, \ell) \equiv (\eta_0 k_{\rm P})^{1-n} \frac{\Gamma(3-n)\Gamma(\ell - \frac{1}{2} + \frac{n}{2})}{2^{3-n}\Gamma^2(2 - \frac{n}{2})\Gamma(\ell + \frac{5}{2} - \frac{n}{2})} , \tag{4.11}$$

where Γ is the gamma function, which for a scale invariant spectrum, $n = 1$, evaluates to

$$f(n = 1, \ell) = \frac{1}{\pi(\ell(\ell+1))} . \tag{4.12}$$

If both the curvature and entropy spectral indexes are close to scale invariant ($n_{\rm s} = n_e = 1$), we find that the so-called *Sachs-Wolfe (SW) plateau* for $\ell \lesssim 20$ is constant:

$$\frac{\ell(\ell+1)}{2\pi} C_\ell = \frac{1}{25}\zeta_0^2 + \frac{4}{25} S_0^2 - \frac{4}{25} \cos(\Delta_c) \zeta_0 S_0 \approx 10^{-10} , \tag{4.13}$$

and the numerical value is the measurement of the DMR instrument aboard the COBE satellite averaged on scales $\lesssim 7^\circ$ (Smoot et al., 1992). Clearly, uncorrelated entropy and curvature perturbations (i.e. with $\cos(\Delta_c) = 0$) both add to the SW plateau, but a positive correlation (defined by $\cos(\Delta_c) > 0$) *reduces* the power on large scales, while a negative correlation increases it, as shown in the top left panel of Fig. 4.9 on page 106. If there is no correlation, the isocurvature Sachs-Wolfe plateau from (4.3) and (4.4) is 36 times larger than the adiabatic one for the same value of Ψ at last scattering, and 4 times larger for the same amplitude of the primordial curvature and entropy perturbations, Eq. (4.13). In the pure adiabatic case, $S_0 = 0$, we obtain from (4.13) an estimate of the primordial amplitude of the curvature perturbation:

$$\zeta_0 \approx 5 \cdot 10^{-5} . \tag{4.14}$$

For models with a non-zero cosmological constant, the Universe becomes Λ dominated for $a/a_0 \geq (\Omega_m/\Omega_\Lambda)^{1/3}$, and the potentials start again to decay. This produces a *late* time ISW which contributes on large scales, where it is dominant with respect to the ordinary SW part described above, producing a rise of the SW plateau at low multipoles. The details differ considerably for adiabatic and isocurvature models, and also depend on the spectral index, see Hu & Sugiyama (1995b) for a detailed explanation.

4.1.2 Acoustic region

The structure of the power spectrum on intermediate scales is the result of several physical effects, sometimes with contrasting impacts. The most distinctive features are acoustic oscillations and projection.

4.1.2.1 Peak locations

Scales $kr_s = k\int_0^{\eta_{\rm dec}} c_s {\rm d}\eta > 1$ enter the horizon before decoupling and thus $D_{g,\gamma}$ oscillates as $\cos(r_s k)$ – cf. (3.56, page 67) – for adiabatic perturbations or as $\sin(r_s k)$ – cf. (3.62, page 67) – in the isocurvature mode. Thus scales which at the moment of decoupling have reached an extremum of their oscillation will yield corresponding peaks in the temperature power spectrum. Notice that since the power spectrum is a quadratic quantity, both maxima and minima of the oscillations give peaks. The k modes which at recombination are at maximum compression or expansion are

$$k_{\rm ad}^{(m)} = \frac{m\pi}{r_s(\eta_{\rm dec})}, \quad m = 1, 2, 3, \ldots \qquad \text{(adiabatic)}, \tag{4.15}$$

$$k_{\rm is}^{(m)} = \frac{m\pi + 1/2}{r_s(\eta_{\rm dec})}, \quad m = 0, 1, 2, \ldots \quad \text{(isocurvature)}. \tag{4.16}$$

The corresponding physical scale $\lambda^{\rm phys} = a_{\rm dec}\pi/k$ subtends an angle ϑ on the sky given by the angular diameter distance relation (1.32, page 27), and the peaks in the angular power spectrum show up at $\ell \sim \pi/\vartheta$ or

$$\ell^{(m)} \sim m\pi \frac{D_A}{a r_s}(\eta_{\rm dec}) \qquad \text{(adiabatic)}, \tag{4.17}$$

$$\ell^{(m)} \sim (\tfrac{1}{2} + m)\pi \frac{D_A}{a r_s}(\eta_{\rm dec}) \quad \text{(isocurvature)}. \tag{4.18}$$

Since $D_{g,\gamma}(k = k_{\rm ad}^{(1)}) < 0$, the first adiabatic peak corresponds to a compression maximum, while the first "isocurvature hump" is an expansion maximum, $D_{g,\gamma}(k = k_{\rm is}^{(0)}) > 0$. In the literature, "first acoustic peak" usually designates the compression peaks, i.e. the first adiabatic extremum and the second isocurvature one, which in the notation of (4.15–4.16) correspond both to the index $m = 1$. For a flat universe ($\mathcal{K} = 0$) without cosmological constant ($\Omega_\Lambda = 0$) and a baryon content as inferred from BBN ($\Omega_b h^2 \approx 0.02$), the location of the first acoustic peak is approximately

$$\ell^{(1)} \sim 220 \quad \text{(adiabatic)} \quad \text{and} \tag{4.19}$$

$$\ell^{(1)} \sim 330 \quad \text{(isocurvature)}. \tag{4.20}$$

The WMAP data allow a very precise determination of the position of the first peak, $\ell^{(1)} = 220.1 \pm 0.8$ (Page et al., 2003), thereby confirming that the adiabatic mode is the dominant one. However, subdominant isocurvature contributions cannot be ruled out, see Chapter 7.

The location of the peaks depends on the of initial conditions, but the inter-peaks distance is independent on the type of perturbations, and in the above estimate is $\Delta\ell \approx 220$. The peak spacing depends on the baryon content, which sets r_s, and on the spatial geometry

which enters in D_A. A larger baryon content slows down the oscillations, thus decreasing the sound horizon and the spacing between peaks grows larger. The dependence of D_A is primarily on the curvature of the universe: in a crude approximation we neglect $\Omega_{\mathcal{K}} \ll \Omega_m$ and Ω_Λ when integrating (1.34, page 27) up to $z_{\rm dec} \approx 1100 \gg 1$ and neglect Ω_r as well (which is not a good approximation for a large redshift) and we obtain

$$D_A(z_{\rm dec}) \approx \frac{2a_{\rm dec}}{H_0 a_0}\Omega_m^{-1/2}\,. \tag{4.21}$$

Therefore the peak position scales as $\Omega_m^{-1/2}$, which means that the peaks are shifted to larger ℓ values in an open universe. Introducing a non-zero cosmological constant complicates matters, since it is then possible to obtain the same value of the angular diameter distance, and hence the same peak location, by compensating a change in Ω_m with a different value of Ω_Λ, an effect which goes under the name of angular diameter distance degeneracy (Efstathiou & Bond, 1999; Melchiorri & Griffiths, 2001). The angular diameter distance test is no longer sufficient to determine alone the curvature of the universe, but an independent measurement of Ω_m or Ω_Λ is necessary.

To illustrate this fundamental degeneracy, let us introduce the *shift parameter* $\mathcal{R}^{\rm shift}$, which gives the first peak's position (in an adiabatic model) with respect to its location in a flat reference model with $\Omega_m = 1$:

$$\ell^{(1)} = \ell^{(1)}_{\rm ref}/\mathcal{R}^{\rm shift}\,, \tag{4.22}$$

which can be evaluated from (4.17). To this end, we need the explicit expression for the sound horizon at decoupling, which is given by

$$\begin{aligned} r_s(a_{\rm dec}) &= \int_0^{a_{\rm dec}} c_s \frac{{\rm d}\eta}{{\rm d}\tilde{a}}{\rm d}\tilde{a} \\ &= \frac{1}{H_0 a_0\sqrt{3}}\int_0^{a_{\rm dec}/a_0} \frac{{\rm d}x}{\left[\left(1+\frac{3\Omega_b}{4\Omega_\gamma}x\right)\left(\Omega_m x + \Omega_r + \Omega_{\mathcal{K}}x^2 + \Omega_\Lambda x^4\right)\right]^{1/2}} \end{aligned} \tag{4.23}$$

(where all the Ω_X's are evaluated today). Neglecting the curvature and cosmological constant term in the early universe ($a_{\rm dec}/a_0 \ll 1$) yields the approximate result

$$\begin{aligned} r_s(a_{\rm dec}) \approx & \frac{1}{\sqrt{3}H_0 a_0 \Omega_m^{1/2}}\left(\frac{a_{\rm eq}/a_0}{R_{\rm eq}}\right)^{1/2} \times \\ & \times \ln\left[\frac{1+R_{\rm eq}+2R_{\rm dec}+2\sqrt{(1+R_{\rm dec})(R_{\rm eq}+R)}}{1+R_{\rm eq}+2\sqrt{R_{\rm eq}}}\right], \end{aligned} \tag{4.24}$$

where

$$R(a) \equiv \frac{3\Omega_b}{4\Omega_\gamma}\frac{a}{a_0} \quad \text{and} \quad R_{\rm eq} \equiv R(a_{\rm eq}), R_{\rm dec} \equiv R(a_{\rm dec})\,. \tag{4.25}$$

In order to find a simple approximate expression for $\mathcal{R}^{\rm shift}$, let us ignore the logarithmic dependence on the parameters of r_s, and neglect the parameter dependence of the factor $(a_{\rm eq}/a_0)^{1/2}/R_{\rm eq}^{1/2}$ as well; we shall relax those approximations in § 4.2. Then the sound horizon for $\mathcal{K} \neq 0$ models scales as

$$r_s(a_{\rm dec}) \approx \alpha\sqrt{\frac{|\Omega_{\mathcal{K}}|}{\Omega_m}}\,, \tag{4.26}$$

while for the reference model with $(\Omega_m, \Omega_\Lambda) = (1, 0)$ we have

$$\frac{D_A(a_{\rm dec})}{a_{\rm dec} r_s(a_{\rm dec})} = 2\alpha\,, \tag{4.27}$$

with α being approximately the same factor as in (4.26). For the shift parameter (4.22) of a model with arbitrary $(\Omega_m, \Omega_\Lambda)$ we then obtain the simple expression

$$\mathcal{R}^{\rm shift} \approx \frac{2}{\chi(\Delta\eta)} \sqrt{\frac{|\Omega_\mathcal{K}|}{\Omega_m}}\,, \tag{4.28}$$

where $\Delta\eta$ is given in Eq. (1.34, page 27) and χ in Eq. (1.3, page 24). This handy expression gives the approximate position of the first peak as a function of Ω_m and Ω_Λ, with $\Omega_\mathcal{K}$ obtained from the constrain $1 = \Omega_m + \Omega_\Lambda + \Omega_\mathcal{K}$. Here we have ignored the dependence on the radiation content of the model, which is explicitly included in (Eq. (6.5, page 138)). In the left panel of Fig. 4.1 we plot lines of $\mathcal{R}^{\rm shift} = $ const in the $(\Omega_m, \Omega_\Lambda)$ plane, which are not parallel to lines of constant curvature (diagonal lines).

Along with $\mathcal{R}^{\rm shift}$, two other physical quantities determine the structure of the peaks: the baryon density $\Omega_b h^2$ controls the relative height of the peaks, see § 4.1.2.2, while the amount of matter $\Omega_m h^2$ sets the redshift of equality, for a fixed relativistic energy content. Therefore by fixing the three quantities $\mathcal{R}^{\rm shift}, \Omega_m h^2, \Omega_b h^2$ we obtain models with almost indistinguishable power spectra in the acoustic region. This is illustrated in the middle panel of Fig. 4.1, where a flat, a closed and an open model result completely degenerate, with the only difference showing up on large scales because of the different amount of late ISW effect. The right panel shows that conversely the first peak's position in three flat models can be very different if the shift parameters differ, and therefore the statement that the first peak position alone can determine the curvature of the Universe is imprecise.

Polarization peaks are displaced by $\pi/2$ with respect to temperature peaks, hence polarization maxima occur at temperature minima. This can be seen by expanding to first order in $\dot\tau^{-1}$ the polarization hierarchy (2.132–2.135, page 54), finding for the polarization monopole and quadrupole

$$\Theta^Q_0 = -\frac{5}{4}\Theta_2 \quad \text{and} \quad \Theta^Q_2 = -\frac{1}{4}\Theta_2\,. \tag{4.29}$$

The temperature quadrupole is found to the same order from the temperature hierarchy, including the polarization feedback as in (2.136, page 54), giving

$$\Theta_2 = -\dot\tau^{-1}\frac{8}{15}\imath k \Theta_1\,. \tag{4.30}$$

The E-polarization source term (3.155, page 84) becomes in the instantaneous decoupling approximation

$$\mathcal{S}^E = -\dot\tau^{-1}(\eta_0 - \eta_{\rm dec})^{-2}\frac{\imath}{k}\Theta_1(\eta_{\rm dec})\,, \tag{4.31}$$

showing that E-polarization probes the temperature dipole, i.e. the bulk velocity of the photons-baryons fluid, at decoupling. Since $\Theta_1 \propto V_\gamma \propto \dot D_{g,\gamma}$ we see that polarization oscillations are out of phase of $\pi/2$, as visible in the top left panel of Fig. 4.9 on page 106.

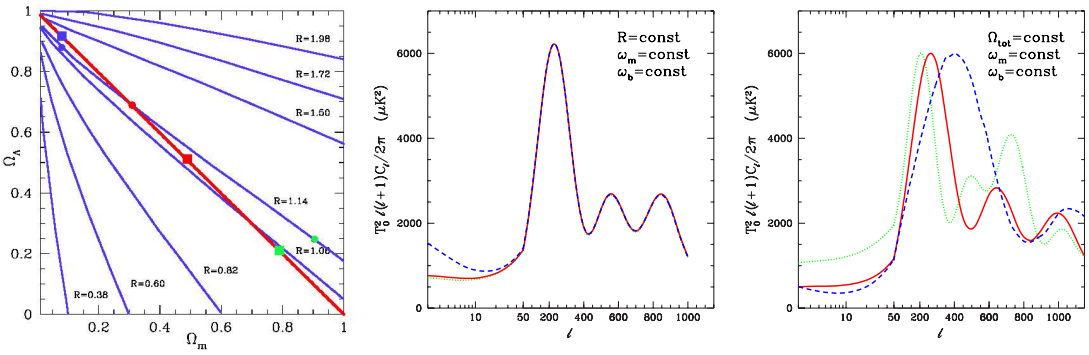

Figure 4.1: Left panel: lines of constant shift parameter (4.28) in the $(\Omega_m, \Omega_\Lambda)$ plane (in blue) correspond to models in which the acoustic peaks are in the same position; those lines are not parallel to lines of constant curvature (in red, the line of $\Omega_{\rm tot} = 1$ is the locus of flat models). Middle panel: a closed (blue, long-dashed), a flat (solid, red) and an open model (dotted green) with parameters corresponding to the three colored dots in the left panel on the $\mathcal{R}^{\rm shift} = 1.14$ line are almost completely degenerate. Right panel: three flat models with different shift parameters (and values corresponding to the three colored squares in the left panel) exhibit a very different peak structure. In particular, measuring the position of the first peak alone is not enough to determine the curvature of the Universe.

4.1.2.2 Baryon signature

Let us now examine in more detail the role of baryons in the adiabatic scenario. The relevant quantity for the final temperature fluctuations is, from Eqs. (3.128) and (3.129, page 80) with $\Phi = \Psi$

$$\begin{aligned}\frac{1}{4}D_{g,\gamma} + 2\Phi - \imath\mu k V_\gamma =& \frac{1}{3}(1+R)\Phi\cos(c_s k\eta) - (2+R)\Phi \\ &+ 2\Phi - \imath\mu\frac{\sqrt{1+R}}{\sqrt{3}}\Phi\sin(c_s k\eta)\,,\end{aligned} \tag{4.32}$$

where we have inserted the adiabatic solution (3.113–3.114, page 77) and explicitly restored the Doppler contribution. The effect of baryons, $R > 0$, is twofold: the amplitude of the cosine oscillation is larger and the zero point is now displaced to $-R\Phi$, i.e. the gravitational effects of falling into and climbing out of the potential at decoupling no longer exactly cancel as in Eq. (4.2), where we had taken $R = 0$. Therefore a larger baryon content enhances compression peaks, which correspond to negative extrema of the cosine[1], while it suppresses expansion peaks. This leads to a distinctive signature of the baryon density on the CMB spectrum: a larger baryon content boosts odd peaks and reduces the even ones, hence a precise measurement of the first three peaks leads to an accurate measurement of the baryon content, as is evident from Fig. 4.6 on page 103.

Up to now we have put aside the Doppler term $V_\gamma \propto \sin(c_s k\eta)$: the sine is out of phase of $\pi/2$ with respect to the density oscillation, and its maxima fill in the zeros of the cosine. In the absence of baryons, this would lead to an exact cancellation and to the disappearance

[1]Note that $\Phi < 0$ inside potential wells, thus $\cos(c_s k\eta) < 0$ indeed gives $D_{g,\gamma} > 0$, according to Eq. (3.113, page 77), i.e. it corresponds to an overdensity with $\delta T/T > 0$.

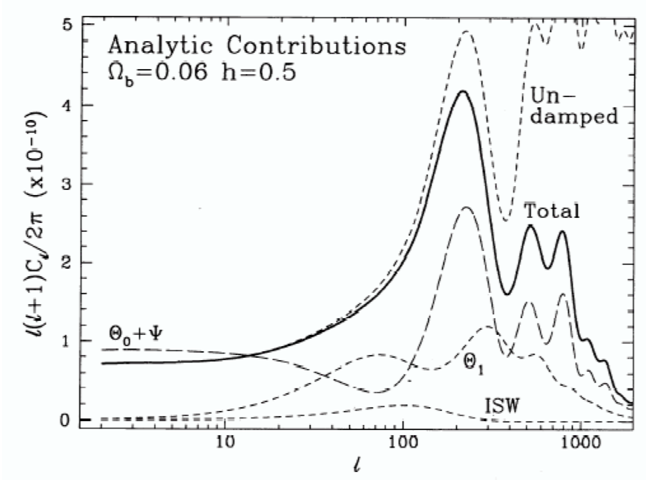

Figure 4.2: Contributions to the adiabatic temperature spectrum (solid) from the temperature monopole (long-dashed), the temperature dipole (Doppler term, short dashed with label Θ_1), and ISW effect (reprinted from Hu & Sugiyama, 1995a).

of the acoustic peaks: adding the density and velocity term incoherently in quadrature for $R = 0$ gives a constant. However, $R > 0$ suppresses the Doppler term by a factor $(1 + R)$ (in quadrature) with respect to the density term, and the net effect is that the velocity contribution partially fills in the minima of the density oscillation without erasing the peak structure, as shown in Fig. 4.2. Also the peak structure for the velocity contribution gets more washed out by the free streaming conversion than for the density, a consequence of the fact that the velocity term is multiplied by μ (Hu & Sugiyama, 1995a).

4.1.2.3 Early ISW effect

At recombination, the Universe is not completely matter dominated, since $a_{\rm dec} \approx 3a_{\rm eq}$ and thus the Bardeen potentials are not exactly constant. This gives an early ISW contribution to the anisotropy, which is spread out over a large multipole range, adding in particular to the rise from the large scale plateau to the first acoustic peak for the adiabatic scenario, cf. Fig. 4.2. Since most of the contribution comes from early times, when $\eta \ll \eta_0$, we can approximatively set $j_\ell(k(\eta_0 - \eta)) \approx j_\ell(k\eta_0)$ and write for the ISW contribution to (3.149, page 83)

$$\Theta_\ell^{(\rm ISW)} = \imath^\ell \int_{\eta_{\rm dec}}^{\eta_0} (\dot\Psi + \dot\Phi) j_\ell(k(\eta_0 - \eta)) \approx \imath^\ell \left[\dot\Psi + \dot\Phi\right]_{\eta_{\rm dec}}^{\eta_0} j_\ell(k\eta_0)\,. \tag{4.33}$$

The early ISW is more prominent if the epoch of equality is delayed due to a smaller matter content or to a larger radiation content, for instance in the presence of extra relativistic particles, as shown in § 6.1.

4.1.3 Damping tail

4.1.3.1 Recombination

Temperature fluctuations on small angular scales are exponentially suppressed by diffusion damping due to the breakdown of tight coupling at recombination, as discussed in § 3.5.

The effect can be roughly incorporated into the undamped solution (3.127, page 80) by multiplying it with the damping factor

$$\mathcal{D}(k) \equiv \int \mathrm{d}\eta g(\eta) e^{-[k/k_{\mathrm{D}}(\eta)]^2} \approx e^{-[k/k_{\mathrm{D}}(\eta_{\mathrm{dec}})]^2}\,, \tag{4.34}$$

using the damping length scale k_{D}^{-1} of Eq. (3.120, page 78).

The main parameter dependence of the damping scale is easy to understand physically: the matter content sets the horizon scale at decoupling, while the baryon density controls the Compton scattering time $\sim \dot{\tau}^{-1}$. Before recombination, photons diffuse by a random walk over a typical length $\lambda_{\mathrm{D}} = \sqrt{N}/\dot{\tau}$, where N is the number of collisions, $N \sim \eta\dot{\tau}$. Hence the damping length scales as

$$\lambda_{\mathrm{D}} \sim \sqrt{\eta_{\mathrm{dec}}/\dot{\tau}} \propto \omega_m^{-1/4}\omega_b^{-1/2}\,, \tag{4.35}$$

where the last proportion takes advantage of the fact that $n_e \propto \omega_b$ (see Eq. (6.17, page 149)) and $\eta_{\mathrm{dec}} \propto \omega_m^{-1/2}$ if decoupling happens in the matter dominated era. A more detailed estimate is given in Eq. (6.19, page 150), which also includes the effect of the helium fraction, which we have ignored here.

Clearly, when recombination occurs the mean free path goes to infinity very rapidly, and therefore the above argument no longer applies, and one has to use a more sophisticated analysis. More details and precise fitting formulas for (4.34) can be found in Hu & White (1997), while useful fitting formulas for many relevant recombination quantities are detailed in Hu & Sugiyama, 1996, Appendix E.

4.1.3.2 Reionization

When the Universe is reionized, the free electron fraction becomes unity again and CMB photons can be rescattered. Fairly little is known about the details of the reionization mechanism and its redshift dependence (for a review see Haiman, 2004) but the null detection of Gunn-Peterson troughs indicates that the Universe was completely ionized after redshift ≈ 6 (Becker et al., 2001), possibly for the second time (Cen, 2003). The recent WMAP results (Spergel et al., 2003) seem to indicate that reionization happened quite early, at a redshift $z_{\mathrm{re}} \approx 17$, corresponding to an optical depth of $\tau_{\mathrm{re}} \approx 0.16$ for a standard ΛCDM model.

Reionization has two effects on the power spectrum: temperature anisotropies on scales below the angle subtended by the horizon at recombination get washed out, and on the same scale there is a generation of polarized power. Let us take for simplicity a model in which all the hydrogen is suddenly reionized at a redshift z_{re}, and ignore helium reionization which happens around $z \approx 3$ which only contributes a few percent. Then the corresponding optical depth to reionization, τ_{re}, is given by

$$\begin{aligned}\tau_{\mathrm{re}} &= \int_{t_0}^{t_{\mathrm{re}}} c\sigma_T n_e \mathrm{d}t \\ &= \frac{c\sigma_T}{H_0}\int_0^{z_{\mathrm{re}}} \frac{n_e(z)}{(1+z)} \frac{\mathrm{d}z}{\left[\Omega_r(1+z)^4 + \Omega_m(1+z)^3 + \Omega_{\mathcal{K}}(1+z)^2 + \Omega_\Lambda\right]^{1/2}}\,.\end{aligned} \tag{4.36}$$

The free electron density (per cm^3) can be expressed as (see Eq. (6.17, page 149))

$$n_e(z) = 11.3 \cdot 10^{-6}(1 - Y_p)\omega_b(1+z)^3\,, \tag{4.37}$$

where we have included the Helium mass fraction Y_p for future reference (see § 6.2.2). For a flat Universe ($\Omega_\mathcal{K} = 0$) and neglecting the contribution of radiation, which is a good approximation if $z_{\rm re} \ll 100$, the integral in (4.36) can be performed analytically, giving (Hu & White, 1997)

$$\tau_{\rm re} = 4.6 \cdot 10^{-2}(1 - Y_p)\frac{\Omega_b h}{\Omega_m}\left[\sqrt{\Omega_\Lambda + \Omega_m(1+z_{\rm re})^3} - 1\right] . \tag{4.38}$$

From the definition of the visibility function g, the probability that a photon last scattered between today and redshift z is

$$P(z) = \int_0^z g(\tilde{z})\mathrm{d}\tilde{z} = 1 - e^{-\tau(z)} , \tag{4.39}$$

and therefore the fraction of photons which arrive to us directly from the recombination epoch is $1 - P(z_{\rm re}) = \exp(-\tau_{\rm re})$. Above the horizon scale at reionization, all photons contribute to the anisotropy, while below that scale only the fraction $\exp(-\tau_{\rm re})$ which did not rescatter contribute. Thus power on small scales will be suppressed by a factor $\exp(-2\tau_{\rm re})$ and the reionization damping factor is given by

$$\mathcal{D}_{\rm re}(k) = \begin{cases} 1 & \text{for } k\tau_{\rm re} \ll 1 \\ e^{-2\tau_{\rm re}} & \text{for } k\tau_{\rm re} \gg 1 \end{cases} . \tag{4.40}$$

The angular scale subtended by the horizon at reionization can be found using (1.32), yielding the approximate scaling (Tegmark & Silk, 1995)

$$\vartheta \propto \sqrt{\frac{\Omega_m}{z}} . \tag{4.41}$$

Without polarization information, reionization is highly degenerate with the spectral tilt and a tensor or isocurvature contribution which would add power only on large scales: a larger reionization optical depth can easily be accommodated by adding tensors or an isocurvature component an reducing at the same time the overall normalization, thereby exactly compensating the reionization power suppression. This degeneracy can be expressed by introducing a suitable combination of $\tau_{\rm re}$ and the overall normalization, see Eq. (4.48) and compare Fig. 4.7. However, the characteristic signature of reionization is the generation of polarized power on the horizon scale of reionization, and the corresponding "polarization bump", clearly visible in the bottom right panel of Fig. 6.16 on page 170, around $\ell \approx 20$ in the E-polarization spectrum can be used to break the degeneracies with other parameters.

The position and scaling of this bump can easily be understood physically (Zaldarriaga, 1997): the temperature quadrupole at reionization, which determines the reionization induced polarization, is given by the free stream of the temperature monopole at decoupling:

$$\Theta_2(\eta_{\rm re}) = (\Theta_0 + 2\Phi)(\eta_{\rm dec}) j_2\left(k(\eta_{\rm re} - \eta_{\rm dec})\right) . \tag{4.42}$$

Given that the k-oscillation of the monopole is much slower than the one of the Bessel function, $r_s \ll \eta_{\rm re} - \eta_{\rm dec}$, the first peak corresponds approximately to the maximum of the Bessel function, which occurs for $k \approx 2/(\eta_{\rm re} - \eta_{\rm dec})$. This translates into $\ell \approx k(\eta_0 - \eta_{\rm re}) \approx 2(\eta_0 - \eta_{\rm re})/(\eta_{\rm re} - \eta_{\rm dec}) \approx 2\sqrt{z_{\rm re}}$. This peculiar scaling of the position of the reionization

bump in the E-spectrum could potentially be used to distinguish the effect of a possible time variation of the fine-structure constant, see § 6.3.4.

Only one parameter is sufficient to characterize the simple model of sudden reionization presented above, namely the reionization redshift $z_{\rm re}$ or equivalently $\tau_{\rm re}$; but it has been shown that there are up to five principal reionization modes which could be extracted from CMB measurements (Hu & Holder, 2003). Furthermore, it is possible to link the reionization history to specific stellar models and try to constrain the parameters of star formation and evolution modelling using CMB data (Bruscoli et al., 2002; Holder et al., 2003; Kaplinghat et al., 2003a).

4.2 Normal parameters

The physical understanding of the characteristic signature of the cosmological parameters can be exploited to build a set of analytical functions which describe quantities directly probed by the CMB. We call such a set a "normal parameter basis", because the effect of the new parameters is almost orthogonal, in the sense that correlations among the parameters should be very small. The normal parameter set has the advantage of taking into account the most severe CMB degeneracies, such as the geometrical degeneracy described above, a feature which improves the efficiency of parameter space exploration (see § 5.1.7). The dependence of the CMB spectrum on the normal parameters is almost linear over a wide range of values, a very important property which makes them ideal as a basis set for the Fisher matrix analysis, see the explanations in § 5.2 and § 6.2.5 for an application. In terms of the normal parameters, it is easy to disentangle and understand the physical effects on the CMB power spectra of each parameter while keeping the other constant, to the contrary of what happens for cosmological parameters.

We have seen in § 4.1.2 that the shift parameter $\mathcal{R}^{\rm shift}$, the baryon and matter density determine the location and relative height of the acoustic peaks. We now expand those considerations by introducing a normal parameter set, based on the discussion of Kosowsky et al. (2002), to which the reader is referred for further details. See also Sandvik et al. (2004) for an application to parameter estimation techniques and Jimenez et al. (2004) for recent improvements including the polarization spectrum.

- The position of the peaks is set by the ratio between the angular diameter distance relation (1.32, page 27) and the physical size of the acoustic horizon at decoupling, Eq. (4.24, page 94). Hence a first normal parameter which determines the overall angular scale is

$$\mathcal{A} \equiv \frac{D_A(a_{\rm dec})}{a_{\rm dec} r_s(a_{\rm dec})}\,, \tag{4.43}$$

  cf. Eq. (4.17), which is just a general expression for the shift parameter. The scale factor at decoupling $a_{\rm dec}$, or equivalently the redshift of decoupling, depends upon $\Omega_b h^2$ and the Ω_m/Ω_r, for which Hu & Sugiyama (1996) provide an accurate analytical fitting formula. The effect of a change in $\mathcal{A}$ while keeping the other normal parameters fixed is displayed in Fig. 4.3.

- The radiation/matter ratio sets the epoch of equality, which in turn determines the

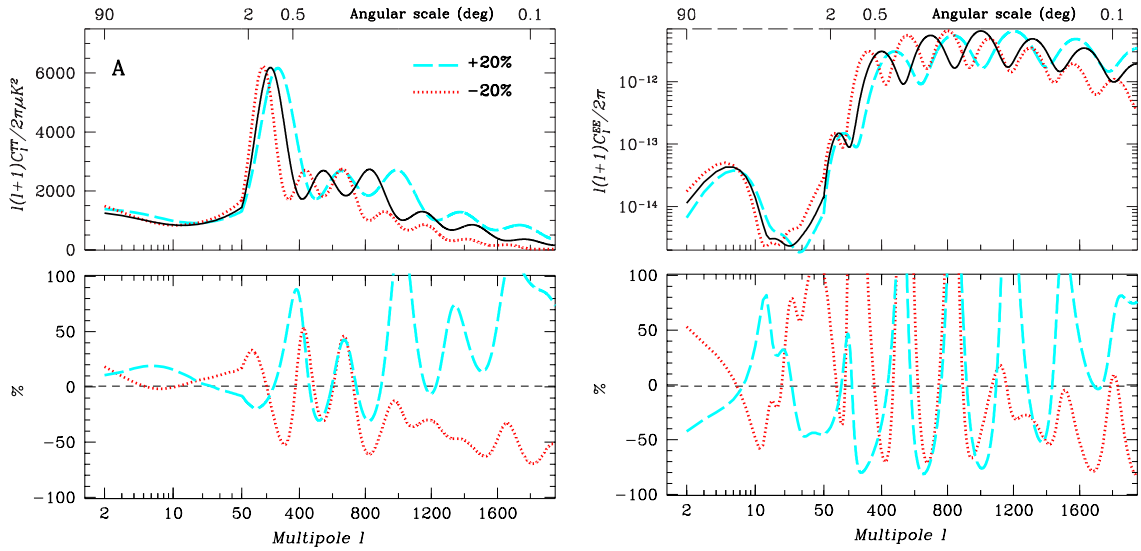

Figure 4.3: Impact of the shift parameter (4.43) on the CMB temperature (left) and polarization (right) spectra, all other normal parameters kept fixed. The geometrical projection effect affects temperature and polarization in the same way. In the bottom panel, we plot the percent difference with respect to the reference model (black).

amount of early ISW, thus we introduce the parameter

$$\mathcal{R} \equiv \frac{\Omega_m}{\Omega_r}\frac{a_{\rm dec}}{a_0}\,, \tag{4.44}$$

which gives the matter to radiation density ratio at the time of decoupling. The boost of the first acoustic peak due to the early ISW is visible in Fig. 4.4.

- The geometrical degeneracy is along the energy density in the cosmological constant, which also gives the amount of late ISW effect. Thus we use the parameter

$$\mathcal{V} \equiv \Omega_\Lambda h^2\,. \tag{4.45}$$

As shown in Fig. 4.5, the impact is quite small in magnitude and solely on large angular scales, where cosmic variance limits our ability to constrain this parameter, making of the cosmological constant one of the worst determinable parameters with CMB data alone.

- The parameter $\mathcal{A}$ already includes the effect of the baryon density on the spacing and location of the peaks, which is produced by the dependence of the sound horizon on the baryon content. Therefore keeping the other normal parameters and in particular $\mathcal{A}$ fixed while varying

$$\mathcal{B} \equiv \Omega_b h^2 \tag{4.46}$$

isolates the baryon driving effect on the acoustic oscillations, which sets the relative height of the peaks. Since the polarization amplitude is proportional to the temperature dipole at recombination, which in turn is suppressed by a factor $(1+R)^{1/2}$ with $R \propto \Omega_b h^2$, a larger baryon density reduces the height of polarization peaks (Fig. 4.6).

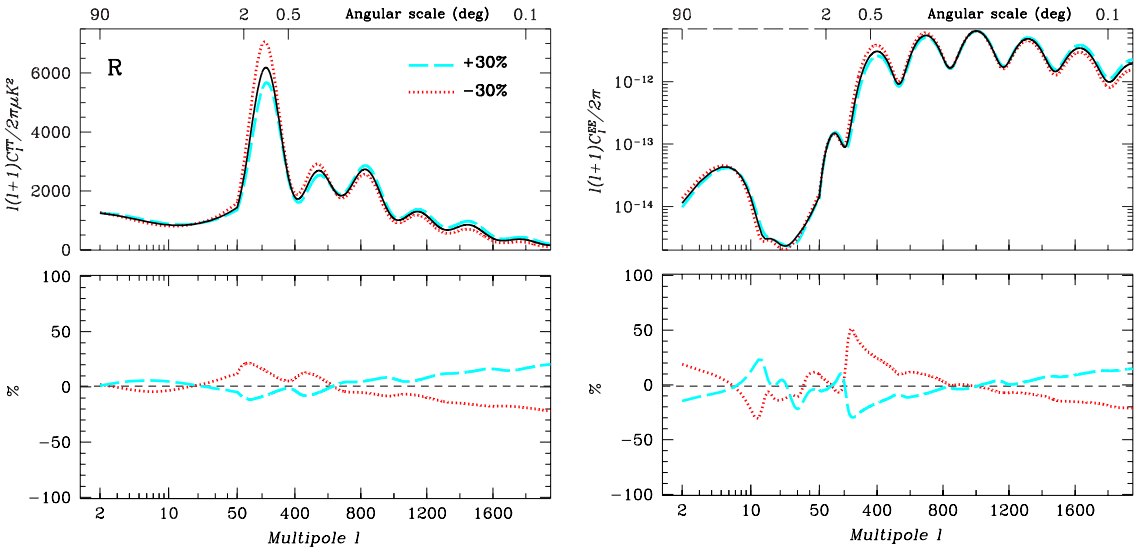

Figure 4.4: Impact of a change in the radiation to matter energy density ratio at decoupling (4.44) on the temperature (left) and polarization (right) spectra, all other normal parameters kept fixed. This can more easily be interpreted as a shift in the epoch of matter-radiation equality, which changes the amount of early ISW effect contribution around the first acoustic peak.

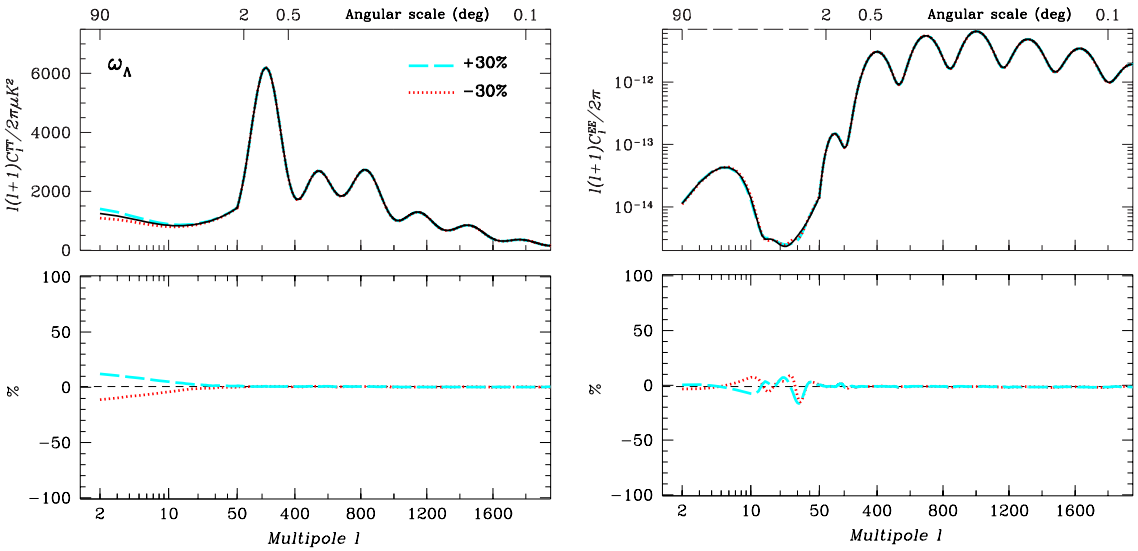

Figure 4.5: Impact of the energy density in the cosmological constant (4.45) on the CMB temperature (left) and polarization (right) spectra, all other normal parameters kept fixed. The impact is only on large angular scales due to the late ISW effect, where measurements are limited by cosmic variance and therefore cannot constraint much this parameter.

- The CMB spectrum turns out to be almost linear in the combination

$$\mathcal{M} \equiv \Omega_m h^2 \left(1 + \frac{\Omega_r^2}{a_{\rm dec}^2 \Omega_m^2}\right)^{1/2} = \Omega_m h^2 \left(1 + \frac{1}{\mathcal{R}}\right)^{1/2}, \tag{4.47}$$

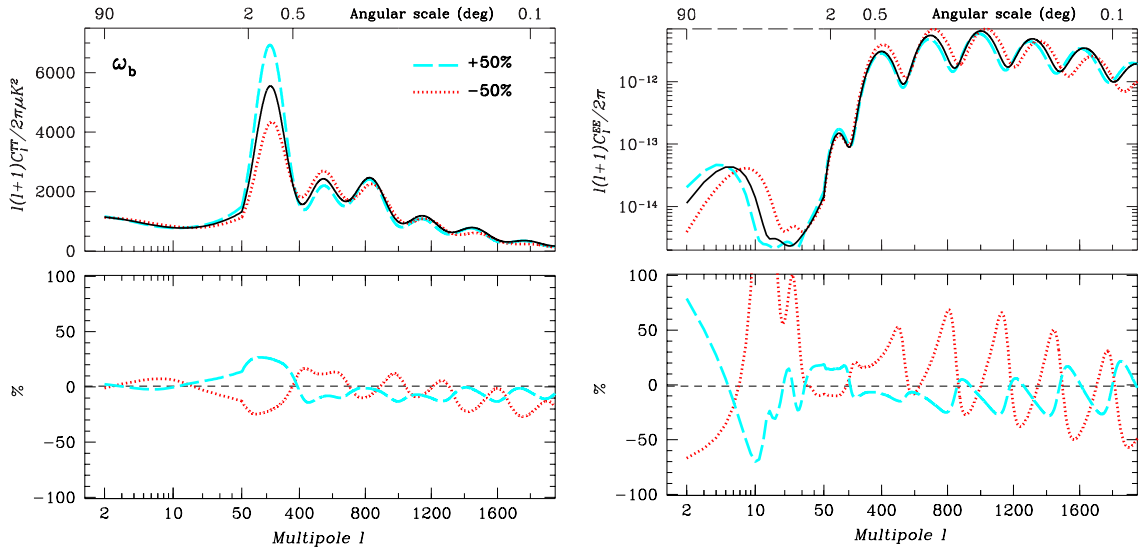

Figure 4.6: Impact of the baryon density (4.46) on the CMB temperature (left) and polarization (right) spectra, all other normal parameters kept fixed. A larger baryon content boosts odd peaks and suppresses even ones, see § 4.1.2.2. The height of the polarization peaks is reduced by a larger baryon content.

which is a refinement of our previous approach of taking simply $\Omega_m h^2$ as a determining parameter, see Kosowsky et al. (2002) for more details.

- A good way of taking into account the degeneracy between the optical depth to reionization and the scalar normalization described in § 4.1.3.2 is to adopt the parameter

$$\mathcal{T} \equiv A_s \exp(-2\tau_{\rm re})\,, \tag{4.48}$$

where for the adiabatic model considered here $A_s \equiv \zeta_0^2$ is the scalar amplitude of the power spectrum of the gauge invariant curvature perturbation, cf. Eq. (4.6, page 91). When adopting a change in $\tau_{\rm re}$, the normalization A_s is also changed as to keep the power above the third peak unchanged, thus avoiding artificial degeneracies with the other normal parameters, which would disappear if one adopted a different normalization convention (Kosowsky et al., 2002), see Fig. 4.7.

- The scale dependence of the initial power spectrum is described by the scalar spectral index n_s, as in (4.6). A value $n_s > 1$ ("blue index") increases the power for wavevectors larger than the pivot scale, and thus yields more power for large multipoles; the converse is true for $n_s < 1$ ("red index"), see Fig. 4.8. Therefore the impact on the CMB spectrum can be approximately modelled as

$$C_{\ell T,E}(n_s) \approx C_{\ell T,E}(n_s = 1) \left(\frac{\ell}{\ell_0}\right)^{n_s - 1} \tag{4.49}$$

with ℓ_0 a pivot point which should be chosen as to match k_P (even though a different choice will only correspond to a change in overall normalization).

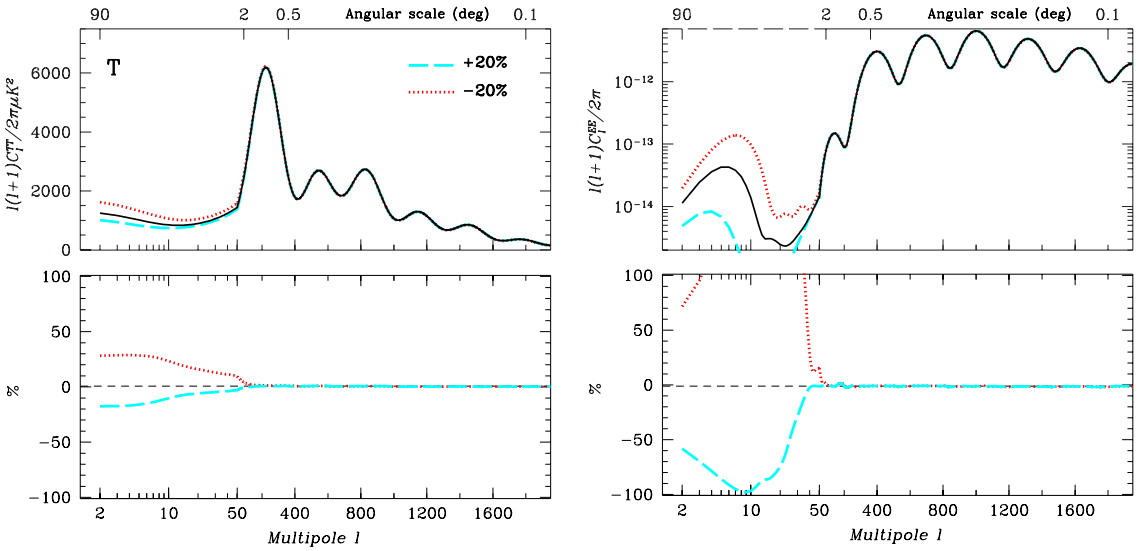

Figure 4.7: Impact of the parameter $\mathcal{T}$ defined in (4.48) on the CMB temperature (left) and polarization (right) spectra, all other normal parameters kept fixed. Increasing $\tau_{\rm re}$ and the overall normalization at the same time as to keep the power above the third peak unchanged reveals the degeneracy between normalization and reionization. The only measurable effect is at large scales, where the temperature signal is *enhanced* for smaller $\mathcal{T}$ (and hence larger $\tau_{\rm re}$) as well as the reionization bump in the polarization spectrum.

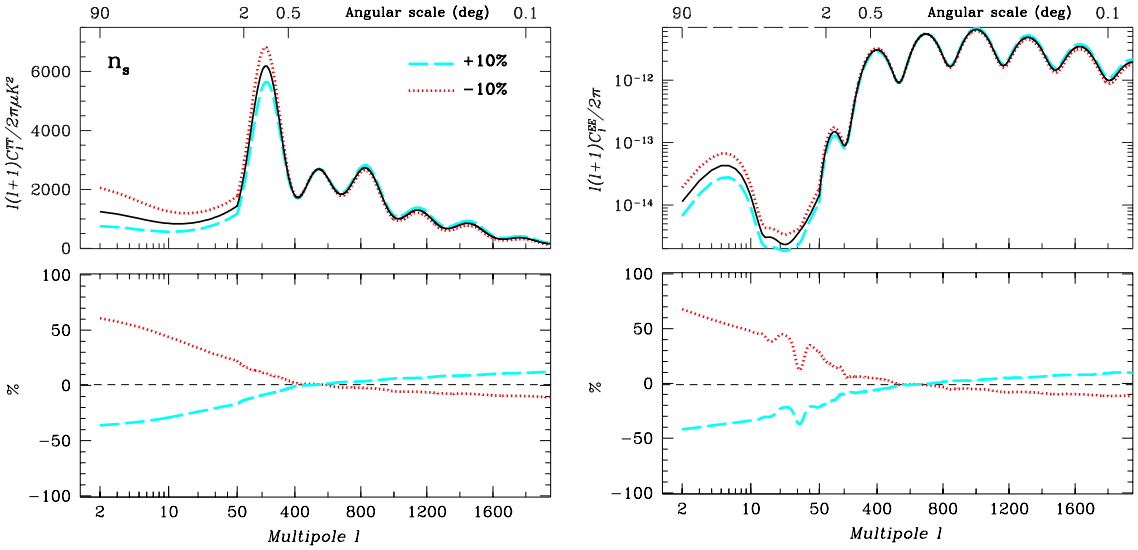

Figure 4.8: Impact of the scalar spectral index on the CMB temperature (left) and polarization (right) spectra, all other normal parameters kept fixed. A blue spectrum ($n_s > 1$) gives more power at larger multipoles. The glitches are numerical artifacts.

Given the above correspondences, we can transform from the cosmological parameter set $(\Omega_m, \Omega_\Lambda, \Omega_b, \Omega_r, h)$ into the normal basis $(\mathcal{A}, \mathcal{R}, \mathcal{V}, \mathcal{B}, \mathcal{M})$ and vice-versa by numerically inverting the relations (4.43–4.47).

4.3 General initial conditions

As we have seen in § 3.2 and § 3.3, a Universe containing photons, massless neutrinos, cold dark matter and photons coupled to baryons admits four growing modes for the perturbations. To this set, one should add a baryon isocurvature entropy mode, which we have not described, but which behaves exactly as the cold dark matter mode, only rescaled by an overall constant $\Omega_b/\Omega_{\rm cdm}$ (Gordon & Lewis, 2003). Thus without loss of generality, we can treat the CDM and baryon isocurvature modes as one single mode, and restrict our considerations to the four modes: adiabatic, CDM isocurvature, neutrino entropy and neutrino velocity.

4.3.1 Angular power spectra for all modes

The numerical integration of the evolution equations is necessary to go beyond the early time approximative solutions derived earlier and obtain the full angular power spectra for the different types of initial conditions. Recent versions of CAMB include the possibility of specifying neutrino entropy and velocity initial conditions, along with the adiabatic and isocurvature CDM ones. The resulting temperature and E-polarization spectra are displayed in Figures 4.9 and 4.10. Analogously to the adiabatic-CDM isocurvature case discussed in § 4.1.1, in the most general case the modes are arbitrarily correlated with each other, and each of them possesses its own spectral index. In the figures we plot the correlators for total positive correlation between the modes, take scale invariant spectral indexes for all modes, $n = 1$ and we fix the other cosmological parameters to a flat, concordance ΛCDM model with early reionization, as emerged from the WMAP data for the pure adiabatic case.

The collection of modes presents a wide variety of oscillatory structures, and very different amplitude ratios between the large-scale plateau and the peaks. Since the perturbation equations are linear, the most general CMB power spectrum is a positive definit superposition of all the modes. From a phenomenological point of view, we expect that widening the initial condition space to include all of the four possible modes, will lead to large degeneracies between initial conditions and cosmological parameters. We dedicate § 7.2 to a thorough investigation of this issue. On the other hand, if the neutrino isocurvature modes were non-zero, their contribution could conceivably allow to fit the CMB data without the need for a cosmological constant, a possibility which we analyze and reject in § 7.3.

4.3.2 Modes superposition

In the purely adiabatic scenario, initial conditions for scalar perturbations are described by two parameters, namely the overall normalization and the spectral index of the curvature perturbation power spectrum, as in Eq. (4.6, page 91). By enlargening the initial conditions space to include all of the four possible modes, we add nine amplitudes (three for the CDM isocurvature, neutrino density and velocity modes, and six for the correlators between the four modes) and three spectral indexes, for a total of 14 parameters describing the most general initial conditions.

Although the dependence of the modes on the amplitudes is trivial, the numerical search in the initial conditions parameter space is complicated by the positive definiteness conditions on the total spectrum. The total temperature (or polarization) angular power spectrum

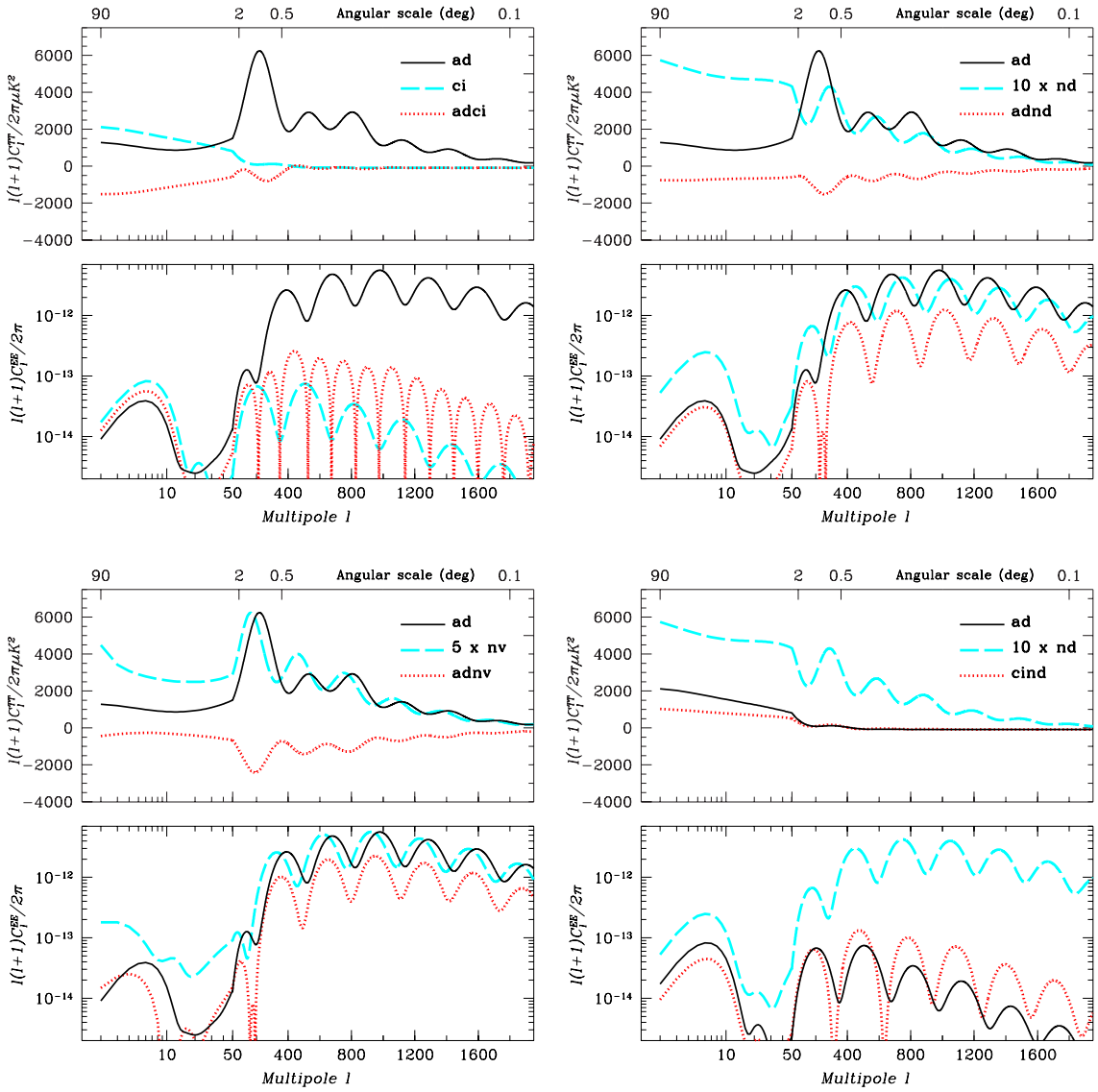

Figure 4.9: Temperature and E-polarization angular power spectra for the four modes constituting the most general initial conditions for CMB anisotropies, Figure 1 of 2. The correlators are for positive total correlation between the modes, and we take all spectral indexes to be unity. The remaining cosmological parameters are fixed to a concordance, flat ΛCDM model. In the lower panel, the correlators are plotted in absolute value. The four modes are: ad (adiabatic), ci (CDM isocurvature), nd (neutrino density/entropy), nv (neutrino velocity).

obtained by superposing the modes must be positive

$$C_\ell = \sum_{i,j=1}^{4} M_{ij} C_\ell^{ij} \geq 0 \quad \forall \quad \ell\,, \tag{4.50}$$

with the *modes correlation matrix* $\mathbf{M} \in \mathcal{P}_n$, where $\mathcal{P}_n$ denotes the space of $n \times n$ real, positive semi-definite, symmetric matrices with in our case $n = 4$, and the C_ℓ^{ij} are computed for a fixed choice of cosmological parameters when only the corresponding element of the

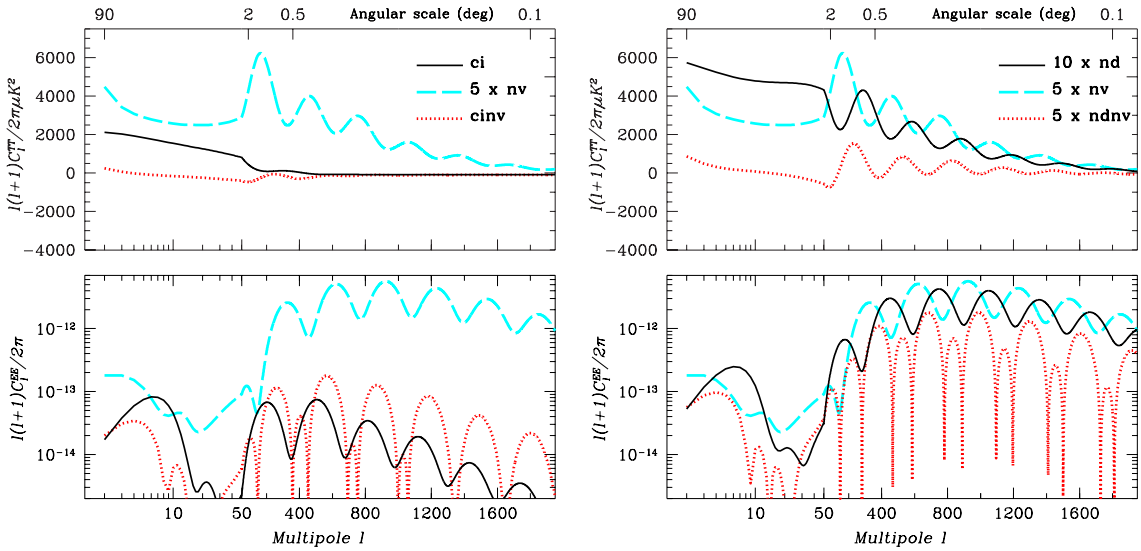

Figure 4.10: Temperature and E-polarization angular power spectra for the four modes constituting the most general initial conditions for CMB anisotropies, Figure 2 of 2.

correlation matrix is non-zero, i.e. for $M_{ij} = 1$, all others vanishing. The elements of the correlation matrix are arranged so that the amplitudes of the pure modes are along the diagonal (so that $M_{ii} \geq 0$ for $i = 1, \ldots, 4$) while the off-diagonal elements are the correlators amplitudes. Each correlator amplitude must satisfy Schwartz' inequality

$$M_{ij}^2 \leq M_{ii} M_{jj} \quad i, j = 1, \ldots, 4 \tag{4.51}$$

because of the positive definiteness condition (see Trotta, 2001, Appendix A for a proof), but in general the correlators amplitudes can of course be negative. Finally, Schwartz' inequality between all pairs $i \neq j$ of $\mathbf{M}$ is a necessary but not sufficient condition for the positive definiteness of the correlation matrix. A sufficient condition is that all sub-determinants of $\mathbf{M}$ are positive or zero (see e.g. Heuser, 1993, proposition 172.5), giving the four sufficient conditions on the elements of $\mathbf{M}$:

$$M_{11} \geq 0\,, \tag{4.52a}$$
$$M_{11} M_{22} - M_{12}^2 \geq 0\,, \tag{4.52b}$$
$$M_{11} M_{22} M_{33} + 2 M_{12} M_{23} M_{13}^2 M_{22} - M_{13}^2 M_{33} - M_{12}^2 M_{33} - M_{23}^2 M_{11} \geq 0\,, \tag{4.52c}$$
$$\det \mathbf{M} \geq 0\,. \tag{4.52d}$$

When numerically searching the initial conditions parameter space, the conditions (4.52) must be imposed by hand to avoid regions which would lead to non-physical (i.e. negative) angular power spectra. This approach is used in Trotta et al. (2001) and some related issues are discussed in § 7.2.

A more convenient parametrization of the correlation matrix is employed in Trotta et al. (2003), where the matrix $\mathbf{M} \in \mathcal{P}_n$ is written as

$$\mathbf{M} = \mathbf{U}\mathbf{D}\mathbf{U}^T, \tag{4.53}$$

$\mathbf{U} \in \mathcal{SO}_n$, $\mathbf{D} = \mathrm{diag}(d_1, d_2, \ldots, d_n)$ and $d_i \geq 0$, $i \in \{1, 2, \ldots, n\}$. Here $\mathcal{SO}_n$ is the space of $n \times n$ real, orthogonal matrices with det = 1 and $n = 4$. We can write $\mathbf{U}$ as an exponentiated linear combination of generators $\mathbf{H}_i$ of $\mathcal{SO}_n$:

$$\mathbf{U} = \exp\left(\sum_{i=1}^{(n^2-n)/2} \alpha_i \mathbf{H}_i\right), \tag{4.54}$$

with

$$\mathbf{H}_1 = \begin{pmatrix} 0 & 1 & 0 & \ldots \\ -1 & 0 & 0 & \ldots \\ 0 & 0 & 0 & \ldots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}, \tag{4.55}$$

and so on, with $-\pi/2 < \alpha_i < \pi/2$, $i \in \{1, 2, \ldots, (n^2 - n)/2\}$. In analogy to the Euler angles in three dimensions, we can re-parameterize U in the form

$$\mathbf{U} = \prod_{i=1}^{(n^2-n)/2} \exp\left(\psi_i \mathbf{H}_i\right), \tag{4.56}$$

with some other coefficients $-\pi/2 < \psi_i < \pi/2$, $i \in \{1, 2, \ldots, (n^2 - n)/2\}$, whose functional relation with the α_i's does not matter. The diagonal matrix $\mathbf{D}$ can be written as

$$\mathbf{D} = \mathrm{diag}\left(\tan(\theta_1), \ldots, \tan(\theta_n)\right), \tag{4.57}$$

with $0 \leq \theta_i < \pi/2$, for $i \in \{1, 2, \ldots, n\}$. In this way, the space of initial conditions for n modes is efficiently parameterized by the $(n^2 + n)/2$ angles θ_i, ψ_j. In our case, $n = 4$ and the initial conditions are described by the ten dimensional hypercube in the variables $(\theta_1, \ldots, \theta_4, \psi_1, \ldots, \psi_6)$. This is of particular importance for the numerical search in the parameter space. One can then go back to the explicit form of $\mathbf{M}$ using Eqs. (4.56), (4.57) and (4.53). This more efficient parametrization is employed in § 7.3.

There is no optimal solution for an efficient and physically motivated parametrization of the initial amplitudes; another possibility, based on a ten-dimensional hypersphere, is employed in the analysis of Bucher et al. (2004).

The fundamental problem of scientific progress, and a fundamental one of everyday life, is that of learning from experience. Knowledge obtained in this way is partly merely description of what we have already observed, but part consists of making inferences from past experience to predict future experience.

Harold Jeffreys
Theory of probability

Part III

PARAMETER EXTRACTION

Chapter 5

Statistics and data analysis

We are now in a position to attack the task of actually determining the values of cosmological parameters from the observed CMB anisotropy. To this end, we need several statistical tools, which we introduce in § 5.1.1. The emphasis is on their application to the CMB: we work out the cosmic variance limit from first principles in § 5.1.2, and we present the Maximum Likelihood principle and its application to data analysis in § 5.1.3; we focus on the differences between the frequentist (§ 5.1.4) and Bayesian approach (§ 5.1.5) to statistics, explaining the procedures to assess likelihood and confidence intervals and their interpretation; we then discuss the implementation of two popular methods to sample the parameters space, the traditional gridding method (§ 5.1.6) and the more efficient Monte Carlo sampling (§ 5.1.7). In § 5.2 we explain the details of the Fisher matrix analysis, an handy and accurate technique to produce forecasts for the expected capabilities in terms of parameters extraction of future CMB observations. In the last section, § 5.3, we offer a brief historical review of the last decade of CMB observation, presenting the data-sets which are then exploited in Chapters 6 and 7.

5.1 Elements of probability and statistics

5.1.1 Some concepts of probability theory

We work in real, three-dimensional space, and we consider a field X which is defined in all points $\mathbf{r} \in \mathbf{R}^3$ in such a way that the probability of obtaining the value X at the point $\mathbf{r}$ is $\mathcal{P}(X, \mathbf{r})$. We call X an *infinite dimensional random field* and $\mathcal{P}$ its 1-point probability distribution function (pdf). In order to fully describe the random field X, we need to specify not only $\mathcal{P}$, but also the 2-point pdf, denoted by $\mathcal{P}_2(X_1, \mathbf{r}_1, X_2, \mathbf{r}_2)$, which describes the probability of getting the value X_1 at the point $\mathbf{r}_1$ and the value X_2 at the point $\mathbf{r}_2$; then the probability distribution for all triples of points, $\mathcal{P}_3$, and so on for an arbitrarily large number of points.

From the definition of probability, the n-point pdf's are not all independent, obeying the relations

$$\mathcal{P}_n(X_1, \ldots, X_n) = \int \mathcal{P}_{n+1}(X_1, \ldots, X_n, X_{n+1}) \mathrm{d}X_{n+1} \,. \tag{5.1}$$

The field X is said to be *statistically homogeneous* if its 1-point pdf is the same in all points

of space:

$$\mathcal{P}(X, \mathbf{r}) = \mathcal{P}(X) \quad \text{(statistical homogeneity)}, \tag{5.2}$$

and *statistically isotropic* if the 2-point pdf depends only on the distance between the points but not on the direction of the vector joining them:

$$\mathcal{P}_2(X_1, \mathbf{r}_1, X_2, \mathbf{r}_2) = \mathcal{P}_2(X_1, X_2, r) \quad \text{(statistical isotropy)}, \tag{5.3}$$

with $r \equiv |\mathbf{r}_1 - \mathbf{r}_2|$. In cosmology, all random fields are assumed to be homogeneous and isotropic. From now on we will always make this assumption. We denote with $\langle\cdot\rangle$ the ensemble average over realizations of the field X (expectation value). For a function $f(X)$, its expectation value is

$$\langle f(X)\rangle \equiv \int_\Omega f(X)\mathcal{P}(X)\mathrm{d}X\,, \tag{5.4}$$

where the integration goes over all possible realizations of X defining the sample space Ω. The expectation value of $f(X) = X$ is called the mean of X. Under the assumption of isotropy, $\langle X\rangle$ is a constant independent on $\mathbf{r}$. Therefore in cosmological perturbation theory we can always take the perturbations to have zero mean, since a constant offset can always be reabsorbed in a redefinition of the corresponding background quantity.

Consider $X(\mathbf{k})$, the harmonic transform of X with respect to the eigenfunctions of the Laplace operator; in $\mathbf{R}^3$ this is the usual Fourier transform. Then as a consequence of homogeneity and isotropy, $X(\mathbf{k})$ has the following properties:

$$\langle X(\mathbf{k})\rangle = \delta^{(\mathrm{D})}(\mathbf{k})\langle X\rangle \tag{5.5}$$

$$\langle X(\mathbf{k}) \cdot X(\mathbf{k}')\rangle = \delta^{(\mathrm{D})}(\mathbf{k} - \mathbf{k}')g(k) \tag{5.6}$$

The *real space correlation function* is defined as

$$\xi(\mathbf{r}) \equiv \langle X(\mathbf{r}_1) \cdot X(\mathbf{r}_1 + \mathbf{r})\rangle\,. \tag{5.7}$$

It is the expectation value of $X_1 \equiv X(\mathbf{r}_1)$ and $X_2 = X(\mathbf{r}_1 + \mathbf{r})$ under the 2-point pdf,

$$\xi(r) = \int \mathrm{d}X_1 \int \mathrm{d}X_2\, \mathcal{P}_2(X_1, X_2, r)X_1X_2\,, \tag{5.8}$$

where in writing $\xi(r)$ instead of $\xi(\mathbf{r})$ we have assumed statistical isotropy.

The field X is called *space ergodic* if we can perform a spatial average instead of an ensemble average and obtain the same result:

$$\lim_{R\to\infty} \left(\frac{4}{3}\pi R^3\right)^{-1} \int_{|\mathbf{r}|<R} f\left[X(\mathbf{r})\right] \mathrm{d}^3\mathbf{r} = \langle f\left[X\right]\rangle\,. \tag{5.9}$$

Notice that ergodicity requires that the field is defined over an infinite space, such as $\mathbf{R}^3$. The temperature field of the CMB however lives on the two-sphere S^2, which is a compact manifolds and therefore not ergodic. Therefore even if we could measure the anisotropies with no experimental error, we still would not be able to perform the ensemble average with perfect accuracy, see § 5.1.2.

We denote by $\hat{f}$ the estimator for $f(X)$, i.e. a procedure applied to a random sample of X to produce a numerical value for f, which is called the estimate. When applied to a

set of observations $X_1^{\rm obs}, X_2^{\rm obs}, \dots X_n^{\rm obs}$ which constitute a random sample, the estimator $\hat{f}$ produces a distribution of estimates, and as such it too is a random variable.

An important particular case is the Gaussian random field, for which all the n-point pdf's are Gaussian. The 1-point pdf is then

$$\mathcal{P}(X) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{X^2}{2\sigma^2}\right), \tag{5.10}$$

while the 2-point pdf is given in terms of the field's correlation function ξ as

$$\mathcal{P}_2(X_1, X_2, r) = \frac{1}{2\pi\sigma^2\sqrt{1-\xi^2(r)}} \exp\left(-\frac{X_1^2 + X_2^2 - 2\xi(r)X_1X_2}{2\sigma^2\left[1-\xi^2(r)\right]}\right) \tag{5.11}$$

and the 2-point pdf (or equivalently, the correlation function) contains the full statistical information.

The statement that the correlation function determines the 2-point pdf completely is true only for a Gaussian field; in general, from (5.8) it is clear that after the integration $\xi(r)$ only contains part of the information encoded in $\mathcal{P}_2$. For instance, Jones (1997) gives an interesting counter-example of a Gaussian and a non-Gaussian distribution with the same correlation function and yet with two different 2-point pdf's.

5.1.2 The origin of cosmic variance

It is instructive to compute explicitly the variance of the observed C_ℓ starting from basic principles. If we assume that the temperature fluctuation Θ is an isotropic and homogeneous random field, then the coefficients of the harmonic expansion on the 2-sphere, the $a_{\ell m}$'s, have zero mean and variance given by the true C_ℓ's:

$$\langle a_{\ell m} \rangle = 0 \tag{5.12}$$

$$\langle a^*_{\ell m} \cdot a_{\ell' m'} \rangle = \delta_{\ell\ell'}\delta_{mm'}C_\ell \,. \tag{5.13}$$

Inflation predicts that the $a_{\ell m}$'s are very close to Gaussian variables, so we make the assumption of Gaussianity and for the pdf of $a_{\ell m}$ we take

$$\mathcal{P}(a_{\ell m}) = \frac{1}{\sqrt{2\pi}C_\ell} e^{-\frac{a_{\ell m}^2}{2C_\ell}} \,. \tag{5.14}$$

The true $a_{\ell m}$'s are of course inaccessible to us, but from the measured temperature fluctuation we obtain an estimate which we denote by $\hat{a}_{\ell m}$. As an estimator for the power spectrum we define

$$\hat{C}_\ell \equiv \frac{1}{2\ell+1} \sum_{m=-\ell}^{\ell} |\hat{a}^2_{\ell m}| = \frac{C_\ell}{2\ell+1} V \,, \tag{5.15}$$

where we have introduced the variable

$$V \equiv \sum_{m=-\ell}^{\ell} \frac{|\hat{a}^2_{\ell m}|}{C_\ell^2} \,. \tag{5.16}$$

Eq. (5.15) implies an ergodic hypothesis, since in the estimator we replaced the expectation value in (5.13) by an average over independent azimutal directions by summing over m.

The variable V is a sum of $2\ell+1$ squared Gaussian variables with unit variance, and therefore (Kendall & Stuart, 1977) its pdf is the chi-square pdf with $2\ell+1=l$ degrees of freedom (dof):

$$\mathcal{P}_{\chi^2_l}(V) = \frac{V^{\frac{l-2}{l}}}{2^{l/2}\Gamma(l/2)} e^{-V/2}\,. \tag{5.17}$$

From this we can write down the pdf for the estimator $\hat{C}_\ell$, which is

$$\mathcal{P}(\hat{C}_\ell) = \frac{l}{C_\ell}\mathcal{P}_{\chi^2_l}\left(\frac{l\hat{C}_\ell}{C_\ell}\right) \tag{5.18}$$

which shows that our estimator is distributed according to a chi-square pdf. For $l\to\infty$ the Central Limit Theorem guarantees that the distribution will become Gaussian, hence

$$\lim_{\ell\to\infty}\hat{C}_\ell = C_\ell \tag{5.19}$$

and the estimator is said to be *consistent*. From (5.18) we can calculate the expectation value of $\hat{C}_\ell$, finding

$$\langle\hat{C}_\ell\rangle = C_\ell \quad \text{(unbiasedness)}, \tag{5.20}$$

and its variance

$$\langle\hat{C}^2_\ell\rangle - \langle\hat{C}_\ell\rangle^2 = \frac{2}{2\ell+1}C^2_\ell \quad \text{(efficiency)}. \tag{5.21}$$

We conclude that the fact that there are only $2\ell+1$ independent directions on the sky for a given multipole ℓ limits the efficiency of our estimator for the power spectrum with variance

$$\frac{\langle\hat{C}^2_\ell\rangle - \langle\hat{C}_\ell\rangle^2}{C_\ell} = \frac{2}{2\ell+1} \quad \text{(cosmic variance)}. \tag{5.22}$$

Despite the fact that cosmic variance is a fundamental statistical limit, an ingenious method to circumvent it and to measure the temperature quadrupole with better than cosmic variance precision has recently been proposed by Skordis & Silk (2004).

5.1.3 The principle of Maximum Likelihood

The estimation problem can be generally stated as follows: starting from a limited number of observations, which constitute a random sample, one wants to reconstruct some properties of the underlying pdf. It is simpler to think of the properties of the pdf as unknown parameters, which we seek to determine. Consider a set of n observations $\mathbf{d} = \left\{d_1^{\rm obs}, d_2^{\rm obs}, \ldots, d_n^{\rm obs}\right\}$ of the variable X and a set of p parameters $\boldsymbol{\theta} = \left\{\theta_1^{\rm obs}, \theta_2^{\rm obs}, \ldots, \theta_p^{\rm obs}\right\}$. The measurements have a conditional probability $\mathcal{P}(d_i|\boldsymbol{\theta})$ to be observed given the value $\boldsymbol{\theta}$ for the parameters. The problem at hand is to estimate the joint conditional probability

$$L(\mathbf{d}|\boldsymbol{\theta}) \equiv \prod_{i=1}^{n}\mathcal{P}(d_i|\boldsymbol{\theta}) \tag{5.23}$$

from the observations $\mathbf{d}$. In the above definition, we thought of L as a function of the random variable X; however, once the observations have been done, we can think of L rather as a function of the unknown parameters $\boldsymbol{\theta}$ for a given value of $\mathbf{d}$ and call it the *likelihood function* (LF).

The maximum likelihood (ML) principle affirms that as an estimate for $\boldsymbol{\theta}$ we should choose the value $\boldsymbol{\theta}^*$ which makes the probability of the actual result obtained, $\mathbf{d}$, as large as it can be, i.e.

$$L(\mathbf{d}|\boldsymbol{\theta}^*) \geq L(\mathbf{d}|\boldsymbol{\theta}) \quad \text{(Maximum Likelihood)} \tag{5.24}$$

for all possible values of $\boldsymbol{\theta}$.

Instead of maximizing the LF, one can minimize the quantity

$$\mathcal{L} \equiv -2 \ln L \,, \tag{5.25}$$

which we will call lognormal LF.

If the pdf is Gaussian, then the ML estimation reduces to the usual least square fit: suppose that the measured $d_i^{\rm obs}$ are independent from each other and Gaussian distributed around their (unknown) true values $d_i(\boldsymbol{\theta})$, with variance given by the experimental error $\sigma_i^{\rm obs}$. Then minimizing $\mathcal{L}$ is equivalent to minimization of the quantity

$$\chi^2(\boldsymbol{\theta}) \equiv \sum_{i=1}^{n} \left(\frac{d_i^{\rm obs} - d_i(\boldsymbol{\theta})}{\sigma_i^{\rm obs}} \right)^2 \,, \tag{5.26}$$

which is called the chi-square.

Applied to the problem of parameter extraction from CMB data, the ML prescription means that, given the measured power spectrum, $C_\ell^{\rm obs}$, with errors σ_ℓ, we have to minimize the value of the chi-square by varying the cosmological parameters of interest. This procedure only gives information about the set of parameters which are the "most probable" to have generated the measurements at hand. However, quantifying the error on our estimate for the parameters is a more subtle business, since it involves dwelling into the exact definition of what probability means. There is a long dispute going on among specialists about the correct interpretation of probability, and some fundamental issues are still unresolved. One can take fundamentally two different point of views on the subject, the orthodox (frequentist) approach or the Bayesian point of view, as we now explain. A good introduction to Bayesian methods and a comparison with the sampling theory approach can be found in Box & Tiao (1973), while Kendall & Stuart (1977) give full details about frequentist theory calculations. Jaynes (2003) is a very enjoyable book, which provides a wider perspective on the logic of science and probability theory. A useful textbook with many stimulating examples of Bayesian inference is MacKay (2003). Frodesen et al. (1979) – written by experimentalists who have used on the field the methods described – is more praxis-oriented, and explains in a practical way the statistical mambo-jumbo.

5.1.4 Orthodox probabilities – Confidence intervals

The orthodox definition of probability – also known as "sampling theory" approach – is based on the empirical repeatability of the experiment, see e.g. Jaynes (2003). If an experiment is performed N times and the outcome A occurs in M of this cases, then the probability of the outcome A is

$$P(A) \equiv \lim_{N\to\infty} \frac{M}{N} \,. \tag{5.27}$$

In the case of continuous variables, the concept of probability is defined as the limiting process (5.27) reached from a finite subdivision in N equiprobable intervals of the sample

space (Kendall & Stuart, 1977, Section 7.11, Vol. 1). The frequentist approach allows the definition and interpretation of *exclusion regions* or *confidence intervals* for the parameters, see below. It is the point of view usually adopted in particle physics, where an experiment can be repeated many times under the same circumstances. It is not very popular in cosmology though, where there is only one particular realization to observe.

Confidence intervals – frequentist

Confidence intervals in the frequentist approach have a straightforward interpretation: consider a random variable X whose pdf depends on the parameter θ which we wish to estimate from a random sample $\left\{x_1^{\rm obs}, x_2^{\rm obs}, \ldots, x_N^{\rm obs}\right\}$ with an estimator $\hat{\theta}$. For instance, one can think of θ as the true mean μ of a normal distribution, and the estimator as the sample mean, $\hat{\mu} = N^{-1}\sum_i x_i^{\rm obs}$.

The estimates are distributed according to some pdf, which we denote by $\mathcal{P}_e$. Then a $100\gamma\%$ *confidence interval* for the estimated parameter $\hat{\theta}$ is the range $[\theta_1;\theta_2]$ such that the probability content for the estimator is γ, i.e.

$$P(\theta_1 < \hat{\theta} < \theta_2) \equiv \int_{\theta_1}^{\theta_2} \mathcal{P}_e {\rm d}\theta = \gamma\,. \tag{5.28}$$

Notice that this is a statement about the probability of our *estimate* $\hat{\theta}$ to lie in a certain range, with the interpretation that, if we would draw the N samples L times under identical circumstances, then the estimates produced by $\hat{\theta}$ fall in the range $[\theta_1;\theta_2]$ γL times. Therefore at this stage we are merely making a statement of the distribution of our estimator. If we want to convert this into a confidence statement for the true value θ, we can say that there is a probability γ that the random interval $[\theta_1;\theta_2]$ will cover the true value θ. In other words, in the long run the limits θ_1 and θ_2 are such that the statement

$$\theta_1 < \theta < \theta_2 \tag{5.29}$$

will be true in $100\gamma\%$ of the cases.

Unfortunately, the above interpretation is unapplicable to cosmology, where we cannot draw new samples at will from the underlying distribution, but we have to content ourselves with the only realization we happen to observe. However, we can still use as an estimator the least-square fit to the observed value, and interpret the result in frequentist's terms.

Consider the least-square fit of (5.26), which applied to the CMB power spectrum is

$$\chi^2(\boldsymbol{\theta}) \equiv \sum_\ell \left(\frac{C_\ell^{\rm obs} - C_\ell(\boldsymbol{\theta})}{\sigma_\ell^{\rm obs}}\right)^2\,, \tag{5.30}$$

where the observed $C_\ell^{\rm obs}$ are estimated using the estimator (5.15): since each term is a sum of $2\ell+1$ Gaussian variables squared (the $\hat{a}_{\ell m}$'s), its distribution becomes Gaussian by virtue of the Central Limit Theorem only for large ℓ. The $\sigma_\ell^{\rm obs}$ are the estimated errors from the observations for each multipole, and $\boldsymbol{\theta}$ is the vector containing the p cosmological parameters of interest. The functional dependence of $C_\ell(\boldsymbol{\theta})$ is given by the underlying theory, which we try to falsify by comparing its predictions with the actual observations.

The least-square estimate for $\boldsymbol{\theta}$ – which is equivalent to the ML estimator for Gaussian variables – is the value $\boldsymbol{\theta}^*$ for which the χ^2 reaches the minimum value χ^{2^*}, which is called

least square estimate. Until this point, the least-square estimation makes no assumptions about the underlying pdf for the variables. To the extent to which the $\hat{C}_\ell$'s can be considered as independent Gaussian variables, then the quantity χ^{2^*} is distributed as a chi-square pdf with $f = n - p$ dof, denoted by $\mathcal{P}_{\chi^2_f}$, see (5.17). Here n is the number of multipoles observed and p the number of fitted parameters.

Under these assumptions, the distribution $\mathcal{P}_{\chi^2_f}$ provides a measure of the *goodness of fit*: assume that a given parameter set $\boldsymbol{\theta}_0$ is the correct one, and that the measured chi-square in our Universe for $\boldsymbol{\theta}_0$ is χ^2_0; then if the measurement would be repeated many times in different realizations, the probability that the outcome will be equal or larger than the true value χ^2_0 is

$$P(\chi^2 > \chi^2_0) = \int_{\chi^2_0}^{\infty} \mathcal{P}_{\chi^2_f}(u)\mathrm{d}u \equiv 1 - \gamma_0\,. \tag{5.31}$$

The interpretation in frequentist terms is straightforward: if some other parameters $\boldsymbol{\theta}_1$ have $\chi^2(\boldsymbol{\theta}_1) = \chi^2_1 \gg \chi^2_0$, the chance that $\boldsymbol{\theta}_1$ is the correct set and we are actually seeing a realization far out in the tail of the distribution is very small.

It now remains to define *confidence intervals* for the parameters basing on the above frequentist interpretation: a $100\gamma\%$ confidence interval encompasses parameters whose measured χ^2 is smaller than the value of corresponding to the quantile[1] of $1-\gamma$ for the distribution $\mathcal{P}_{\chi^2_f}$. In other words, if the measurements could be repeated many times, in the long run the above confidence interval would include the true value of the parameters $100\gamma\%$ of the time. Thus the parameter space outside the estimated confidence interval is a proper *exclusion region* at the given confidence level. Notice that the frequentist confidence levels depend both on the *total* number of parameters fitted and on the number of independent data points we are using.

We conclude this section with two remarks: firstly, the above assumptions of Gaussianity and independency are only partially fulfilled by the $\hat{C}_\ell$'s, therefore the outcome of such a frequentist analysis is only approximative (see Abroe et al., 2002 for a strictly correct frequentist parameter estimation, which involves the numerical sampling of the pdf which we simply took as a chi-square); and second, the clean interpretation of the frequentist approach is somewhat weakened by the fact that we are compelled to invoke measurements in other realizations which cannot take place, not even in principle. Bayesian statistics takes instead a more pragmatic approach, by dealing only with actual observations.

5.1.5 Statistical inference – Likelihood intervals

Bayesian statistics does not consider possible outcomes of measurements which are never performed. Instead, it exploits the actual data to update our knowledge about the probability of a certain statement, starting from our prior degree of belief. Criticism has been raised against this approach because the final inference depends on the prior information available, and therefore seems to suffer from a certain degree of subjectivity. However, Bayesian inference can be applied to theories which are not repeatable and are unscientific in the frequentist point of view (e.g. the probability that it will rain tomorrow). It is based on Bayes'

[1] Given the pdf $\mathcal{P}$, x is said to be the quantile of q if it satisfies $\int_x^\infty \mathcal{P}(u)\mathrm{d}u = q$.

Theorem[2], which is nothing more than rewriting the definitions of conditional probability:

$$\mathcal{P}(A|B) = \frac{\mathcal{P}(B|A)\mathcal{P}(A)}{\mathcal{P}(B)} \quad \text{(Bayes' Theorem)}. \tag{5.32}$$

In order to clarify the meaning of this relation, let us write $\boldsymbol{\theta}$ for A and $\mathbf{d}$ for B, obtaining

$$\mathcal{P}(\boldsymbol{\theta}|\mathbf{d}) = \frac{L(\mathbf{d}|\boldsymbol{\theta})\mathcal{P}(\boldsymbol{\theta})}{\int \mathrm{d}\boldsymbol{\theta}\mathcal{P}(\mathbf{d}|\boldsymbol{\theta})\mathcal{P}(\boldsymbol{\theta})} = \frac{L(\mathbf{d}|\boldsymbol{\theta})\mathcal{P}(\boldsymbol{\theta})}{\mathcal{P}(\mathbf{d})}\,, \tag{5.33}$$

which relates the *posterior probability* $\mathcal{P}(\boldsymbol{\theta}|\mathbf{d})$ for the parameters $\boldsymbol{\theta}$ given the data $\mathbf{d}$ to the likelihood function $L(\mathbf{d}|\boldsymbol{\theta})$ if the *prior pdf* $\mathcal{P}(\boldsymbol{\theta})$ for the parameters is known. The quantity in the denominator is independent of $\boldsymbol{\theta}$ and it is called the *evidence* of the data for a certain model (MacKay, 2003). It is important for model comparison, but here we shall regard it just as a normalization constant. In short

$$\text{posterior} = \frac{\text{likelihood} \times \text{ prior}}{\text{evidence}}\,. \tag{5.34}$$

The prior distribution contains all the (subjective) knowledge about the parameters before observing the data: our physical understanding of the model, our insight into the experimental setup and its performance, in short the amount of all our prior scientific experience. This information is then updated via Bayes theorem to the posterior distribution, by multiplying the prior with the LF which contains the information coming from the data. The posterior probability is the base for inference about $\boldsymbol{\theta}$: the most probable value for the parameters is the one for which the posterior probability is largest.

Bayes' postulate[3] states that in absence of other arguments, the prior probability should be assumed to be equal for all values of the parameters over a certain range, $\boldsymbol{\theta}_{\min} \leq \boldsymbol{\theta} \leq \boldsymbol{\theta}_{\max}$. This is called a "flat prior", i.e.

$$\mathcal{P}(\boldsymbol{\theta}) = [H(\boldsymbol{\theta} - \boldsymbol{\theta}_{\min})H(\boldsymbol{\theta}_{\max} - \boldsymbol{\theta})] \prod_{i=1}^{p} [\theta_{\max,i} - \theta_{\min,i}]^{-1}\,, \tag{5.35}$$

where H is the Heaviside step function and $\boldsymbol{\theta}_{\max,i} > \boldsymbol{\theta}_{\min,i}\ \forall\ i$. This is one of the principal conceptual difficulties of Bayesian inference: a flat prior on $\boldsymbol{\theta}$ does not correspond to a flat prior on some other set $f(\boldsymbol{\theta})$, obtained via a non-linear transformation f. Therefore the result of Bayesian inference do depend on the choice of priors, even though this usually does not constitue a major obstacle in practical problems – see however Bucher et al. (2004) for an instructive example of the role of priors.

We see from Eq. (5.33) that the Maximum Likelihood principle is equivalent to Bayesian inference in the case of flat priors. We will always work with flat, top-hat priors unless otherwise stated. There is however an important conceptual difference. By writing the posterior distribution as

$$\mathcal{P}(\boldsymbol{\theta}|\mathbf{d}) = \frac{\mathcal{P}(\boldsymbol{\theta}, \mathbf{d})}{\mathcal{P}(\mathbf{d})}\,, \tag{5.36}$$

it follows that Bayes' Theorem imposes to maximise the *joint probability* $\mathcal{P}(\boldsymbol{\theta}, \mathbf{d})$ of $\boldsymbol{\theta}, \mathbf{d}$, while Maximum Likelihood requires that the *conditional probability* $L(\mathbf{d}|\theta)$ should be maximised.

[2]Rev. Thomas Bayes, 1763.

[3]Bayes' postulate is also known – perhaps with an hint of sarcasm – as the Postulate of Equidistribution of Ignorance.

Likelihood intervals – Bayesian

Bayesian statistics use the LF to perform an interval estimation for $\boldsymbol{\theta}$: basing on Bayes' Theorem, Eq. (5.33), we not only consider the ML point in parameter space as the "most likely" value of the unknown parameter; we shall also interpret values further and further away as less and less likely to have generated the particular measurement at hand. Hence *likelihood intervals* drawn from the LF measure our "degree of belief" that the particular set of observations was generated by a parameter belonging to the estimated interval. This is radically different from the frequentist interpretation sketched above.

Let us simplify the notation by writing $L(\boldsymbol{\theta})$ instead of $L(\mathbf{d}|\boldsymbol{\theta})$, since now we consider the LF as a function of the parameters given a data set $\mathbf{d}$. Assume further that the LF is a multivariate Gaussian distribution in the p parameters $\boldsymbol{\theta}$, i.e.

$$L(\boldsymbol{\theta}) = (\det \mathbf{C})^{-1/2}(2\pi)^{-p/2}\exp(-\mathcal{L}/2)\,, \tag{5.37}$$

$$\mathcal{L} = -2\ln L = (\boldsymbol{\theta}-\boldsymbol{\mu})^T\mathbf{C}^{-1}(\boldsymbol{\theta}-\boldsymbol{\mu}) \tag{5.38}$$

where T denotes transposition, $\boldsymbol{\mu}$ is the expectation value of the parameters $\boldsymbol{\mu} \equiv \langle\boldsymbol{\theta}\rangle$ and $\mathbf{C}$ is the *covariance matrix*

$$C_{ij} \equiv \langle(\theta_i-\mu_i)(\theta_j-\mu_j)\rangle\,. \tag{5.39}$$

From the likelihood one can then obtain the posterior distribution via (5.33), once the prior is specified. For the prior distribution $\mathcal{P}(\theta)$ a simple choice are so-called "flat" priors, a multidimensional top-hat function over some range which is supposed to encompass all the values of interest. Usually, in grid-based method the prior coincides with the extension of the grid, so that the prior is just a multiplicative constant and we can identify the likelihood with the posterior. As mentioned, this choice is somewhat arbitrary, since it depends on the basis chosen for the parameters.

We can Taylor expand a general LF around its maximum which is given by our ML estimate $\boldsymbol{\theta}^*$ of $\boldsymbol{\mu}$, which on average coincides with the true mean for a normal distribution, $\langle\boldsymbol{\theta}^*\rangle = \boldsymbol{\mu}$. By definition of the ML point the first derivatives vanish, $\partial\mathcal{L}/\partial\theta_i(\boldsymbol{\theta}^*) = 0$, and we obtain

$$\mathcal{L}(\boldsymbol{\theta}) \approx \mathcal{L}(\boldsymbol{\theta}^*) + \frac{1}{2}\sum_{ij}(\theta_i-\theta_i^*)\frac{\partial^2\mathcal{L}}{\partial\theta_i\partial\theta_j}(\theta_j-\theta_j^*)\,. \tag{5.40}$$

If the LF is sharply peaked around $\boldsymbol{\theta}^*$, i.e. the errors on the parameters are small enough, then third order terms are unimportant and the above Gaussian form is a good enough approximation everywhere in parameter space. By comparing with (5.38) we find that the covariance matrix can thus be estimated as

$$\hat{\mathbf{C}} = \mathbf{F}^{-1} \quad \text{where} \quad F_{ij} \equiv \left\langle\frac{1}{2}\frac{\partial^2\mathcal{L}}{\partial\theta_i\partial\theta_j}\right\rangle\Big|_{\boldsymbol{\theta}^*} \tag{5.41}$$

is called *Fisher information matrix* (Kendall & Stuart, 1977, Chap.15, Vol.1).

According to our understanding of the LF as a measure of our degree of belief for the possible values of $\boldsymbol{\theta}$, the probability that parameters within a certain region from the ML point have generated the observations should be proportional to the likelihood content of the region. The probability content depends on whether we are estimating all parameters jointly, or keeping some of them fixed to their ML value, or rather disregarding a certain subset by integrating over them (marginalization). We consider each case in turn.

Estimation of all p parameters jointly.

Without loss of generality we can take in the following $\boldsymbol{\mu} = \mathbf{0}$ in Eq. (5.38), which can always be achieved by shifting the origin of the coordinate system in parameter space. Contours of constant likelihood define hyperellipses in parameter space with some probability content we wish to determine. To this aim we consider the quadratic form

$$Q(\boldsymbol{\theta}) \equiv \boldsymbol{\theta}^T \mathbf{C}^{-1} \boldsymbol{\theta} \tag{5.42}$$

and for the LF (5.37) the condition $Q(\boldsymbol{\theta}) = Q^s_\gamma$ for some constant Q^s_γ gives the contours of constant likelihood. We write Q^s_γ to indicate that the numerical value of the constant depends on the number of parameters under consideration, s, and on the desired probability content of the hyperellipse, γ. It can be shown (Kendall & Stuart, 1977, Chap.8, Vol.1) that the quadratic form Q is chi-square distributed with s dof, which allows us to relate Q^s_γ with the probability content of the ellipse.

If we want a confidence region containing $100\gamma\%$ of the *joint* probability for all p parameters, then $s = p$ and Q^p_γ is determined by solving

$$\int_0^{Q^p_\gamma} \mathcal{P}_{\chi^2_p}(u) \mathrm{d}u = \gamma \,. \tag{5.43}$$

The projection (*not* the intersection) of the hyperellipse $Q(\boldsymbol{\theta}) = Q^p_\gamma$ onto each of the parameter axis gives the corresponding likelihood interval for each parameter when all parameter are estimated simultaneously (which we will call "joint likelihood interval").

It is a simple geometrical problem to find an analytical expression for the joint likelihood interval for each parameter: for the parameter $1 \leq d \leq p$, the intersection of the hyperellipse with the hyperplane defined by $\theta_d = c$, with c a constant, gives either an hyperellipse in $p-1$ dimensions, or a point or else an empty set. The extrema of the joint likelihood interval for the parameter d are given by the values of c for which the $p-1$ dimensional ellipse reduces to a point.

To find the equation of the $p-1$ dimensional ellipse we proceed as follows: define $\mathbf{C}^{-1} \equiv \mathbf{M}$ and write $Q(\boldsymbol{\theta}) = Q^p_\gamma$ in the form

$$\tilde{\boldsymbol{\theta}}^T \tilde{\mathbf{M}} \tilde{\boldsymbol{\theta}} + 2c \sum_{j \neq d} m_{dj} \tilde{\theta}_j = Q^p_\gamma - m_{dd} c^2 \,, \tag{5.44}$$

where we have defined

$$\tilde{\boldsymbol{\theta}} \equiv (\theta_1, \ldots, \theta_{d-1}, \theta_{d+1}, \ldots, \theta_p) \in \mathbf{R}^{p-1} \tag{5.45}$$

$$\tilde{\mathbf{M}} \equiv \begin{pmatrix} m_{11} & \ldots & m_{1,d-1} & m_{1,d+1} & \ldots & m_{1p} \\ \vdots & & & & & \vdots \\ m_{d-1,1} & & \ldots & & & m_{d-1,1} \\ m_{d+1,1} & & \ldots & & & m_{d+1,1} \\ \vdots & & & & & \vdots \\ m_{p1} & & \ldots & & & m_{pp} \end{pmatrix} \in \mathbf{R}^{(p-1 \times p-1)}. \tag{5.46}$$

Now we diagonalize the submatrix $\tilde{\mathbf{M}}$,

$$\mathrm{diag}\,(\lambda_1, \ldots, \lambda_{p-1}) \equiv \boldsymbol{\Lambda} = \mathbf{U}^T \tilde{\mathbf{M}} \mathbf{U} \tag{5.47}$$

finding the eigenvalues $\lambda_i, i \leq 1 \leq p-1$ and eigenvectors $(u_1, \ldots, u_{p-1})$, and after some algebraic manipulations of (5.44) we arrive at the equation of the $p-1$ dimensional hyperellipse

$$\sum_{i=1}^{p-1} \lambda_i z_i^2 = Q_\gamma^p - m_{dd}c^2 + \sum_{i=1}^{p-1} \frac{c^2}{\lambda_i} \left(\sum_{j \neq d} m_{dj} u_{ji} \right)^2 , \tag{5.48}$$

where we have defined the new variables

$$z_i \equiv (\tilde{\theta}\tilde{\mathbf{U}})_i + \frac{c}{\lambda_i} \sum_{j \neq d} m_{dj} u_{ji} , \quad 0 \leq i \leq p-1 . \tag{5.49}$$

The above hyperellipse becomes degenerate if

$$\sum_{i=1}^{p-1} \lambda_i z_i^2 = 0 \tag{5.50}$$

from which we obtain a quadratic equation for c with solutions

$$c_{\text{min, max}} = \frac{\pm\sqrt{Q_\gamma^p}}{\left[m_{dd} - \sum_{i=1}^{p-1} \lambda_i^{-1} \left(\sum_{j \neq d} m_{dj} u_{ji} \right)^2 \right]^{1/2}} . \tag{5.51}$$

It is easy to show that the positive definiteness condition for the Fisher matrix guarantees that the quantity under the square root in the denominator is always ≥ 0. In conclusion, the joint likelihood interval for the parameter θ_d with likelihood content γ is given by

$$c_{\min} \leq \theta_d \leq c_{\max} . \tag{5.52}$$

Estimation of $k < p$ parameters, the others fixed.

We are sometimes interested in giving confidence intervals for some subset $k < p$ of the parameters, while assuming the other $p-k$ parameters as (exactly) known. Without loss of generality we shall take the first k parameters as the one we are interested in, and we split the parameter vector as

$$\theta = \begin{pmatrix} \mathbf{t} \\ \mathbf{u} \end{pmatrix} \tag{5.53}$$

with $\mathbf{t} \in \mathbf{R}^k$ and $\mathbf{u} \in \mathbf{R}^{p-k}$. Correspondingly we write the covariance matrix in (5.38) as the Fisher matrix estimate of (5.41),

$$\mathbf{C}^{-1} = \mathbf{F} = \begin{pmatrix} \mathbf{A} & \mathbf{G} \\ \mathbf{G}^T & \mathbf{B} \end{pmatrix} \tag{5.54}$$

where $\mathbf{A} \in \mathbf{R}^{k \times k}$, $\mathbf{B} \in \mathbf{R}^{p-k \times p-k}$ and $\mathbf{G} \in \mathbf{R}^{p-k \times k}$.

If the known parameters $\mathbf{u}$ are held fixed at their ML value, the LF for the parameters of interests $\mathbf{t}$ is simply the full LF restricted to the k subspace,

$$L(\mathbf{t}|\mathbf{u}^*) \propto \exp(-\frac{1}{2}\mathbf{t}^T\mathbf{A}\mathbf{t}) , \tag{5.55}$$

with an appropriate normalization constant, and the new covariance matrix $\mathbf{V} \in \mathbf{R}^{k\times k}$ for the k parameters of interest is

$$\mathbf{V} = \mathbf{A}^{-1} \quad \text{(conditional)}. \tag{5.56}$$

In particular, we often consider the best case scenario in which all parameters but one are supposed to be known exactly, say from independent observations or theoretical prejudice, and therefore $k = 1$. Then the 1σ likelihood interval for the first parameter only is the square root of the covariance matrix element, and it is given by (all others fixed to their ML value)

$$\sigma_1 = \frac{1}{\sqrt{f_{11}}}\,. \tag{5.57}$$

Estimation of $k < p$ parameters, the others marginalized.

Instead of fixing some parameters, we may prefer to disregard them completely, by integrating over them in order to obtain the *marginalized* likelihood in the k parameter of interest:

$$L(\mathbf{t}) \propto \int_{\Omega_\mathbf{u}} L(\mathbf{t},\mathbf{u})\mathrm{d}\mathbf{u}\,, \tag{5.58}$$

with a suitable normalization constant so that the probability content of the marginalized LF is equal to unity.

The marginal LF for $\mathbf{t}$ is still a multivariate Gaussian, with the same covariance matrix as the full LF, only with the last $p - k$ rows and columns deleted:

$$V_{ij} = \left[\mathbf{F}^{-1}\right]_{ij} \quad 0 \le i,j \le k \quad \text{(marginalized)}. \tag{5.59}$$

This result can be obtained by performing explicitly the integration (5.58) or more elegantly by using the properties of the characteristic function (Kendall & Stuart, 1977, Chap.4, Vol.1). In terms of the splitting (5.54), the covariance matrix for the marginalized distribution is

$$\mathbf{V} = \left[\mathbf{A} - \mathbf{G}\mathbf{B}^{-1}\mathbf{G}^T\right]^{-1}\,. \tag{5.60}$$

Very often one quotes marginalized likelihood intervals for one parameter alone, $k = 1$ with all other parameters marginalized, in which case the 1σ error is given by

$$\sigma_1 = \sqrt{(\mathbf{F}^{-1})_{11}}\,. \tag{5.61}$$

If the parameters are uncorrelated, then $\mathbf{F}$ is diagonal, and fixing $\mathbf{u}$ or marginalizing over them is equivalent, otherwise the resulting likelihood intervals for the parameter(s) of interest are in general different, with the marginalized interval being broader.

5.1.6 Gridding method

In the numerical fit to the data, the shape of the LF is determined by evaluating the least-square estimator (5.26, page 115) at each point on a grid in the p dimensional parameter space and the minimization of the chi-square in the desired range of parameters gives the ML estimate.

$100\gamma\%$ Likelihood content	68.3% (1σ)	95% (1.96σ)	95.4% (2σ)	99% (2.58σ)	99.7% (3σ)
1 parameter, Q^1_γ	1.00	3.84	4.00	6.63	9.00
2 parameters, Q^2_γ	2.30	5.99	6.17	9.21	11.80

Table 5.1: $\Delta\chi^2 = Q^k_\gamma$ for marginalized likelihood intervals in one parameter ($k = 1$) or marginalized likelihood contours in two parameters ($k = 2$) for the given joint likelihood content.

Assuming that the measurements are normally distributed around their true value we have

$$L(\mathbf{d}|\boldsymbol{\theta}) = L_{\max} \exp\left[-\chi^2(\boldsymbol{\theta})/2\right] . \tag{5.62}$$

From this we can use the above prescriptions to determine likelihood or confidence intervals from real data.

In the frequentist analysis, the boundaries of the confidence regions represent exclusion plots at the given confidence level: they are found as the contours of constant χ^2 using the relation (5.31, page 117), independently of the value of the chi-square at the ML point. In Bayesian statistics, the likelihood intervals are instead drawn *around* the ML point, hence their extension depends on the best fit value. This applies only to the gridding method, not to the Monte Carlo sampling described below in § 5.1.7. It is customarily to quote marginalized likelihood intervals for one parameter only or to plot two-dimensional likelihood contours to show degenerate direction between two parameters (also see below the paragraph discussing the maximization approach instead of marginalization); for these two cases, the cook-book prescription for Bayesian (Maximum Likelihood) statistics on a grid of samples in parameter space is:

- find the ML point $L_{\max}$ in the grid of parameters by minimizing the χ^2 of Eq. (5.30, page 116) and mark this point as $\chi^2_{\min}$, your least-square estimate of the best fit;
- determine the boundaries of the region containing $100\gamma\%$ of likelihood as the values of the parameters for which the χ^2 has increased by an amount $\Delta\chi^2 = Q^k_\gamma$ ($k = 1, 2$ the number of parameters considered) with respect to $\chi^2_{\min}$.
- The values of Q^k_γ can be found for every desired likelihood content using the relation, cf. (5.31, page 117)

$$\gamma = \int_0^{Q^k_\gamma} \mathcal{P}_{\chi^2_k}(u)\mathrm{d}u\,. \tag{5.63}$$

  Table 5.1 displays the values of $\Delta\chi^2$ for $k = 1, 2$ and for some popular choices of likelihood content.

In a real situation, the LF computed using (5.62) will not be exactly a multivariate Gaussian, and the likelihood intervals obtained with this method will only approximatively encompasses the stated probability content. There are methods which improve on the assumption of a normal distribution presented here, see for instance Bond et al. (2000); Bartlett et al. (2000); Wandelt et al. (2001); Jaffe et al. (2003).

Finally, notice that likelihood (Bayesian) contours are usually much tighter than the confidence contours drawn from the frequentist point of view. This is a consequence of the ML point having often a χ^2/f much smaller than 1, because the data-sets are highly consistent with each other and also because usually not all points are completely independent. For the CMB, this was the case when one considered a combination of several data-sets before WMAP, as we discuss in § 7.2. If we consider the usual situation in which likelihood contours are drawn in a two dimensional plane with all other parameters marginalized over, the frequentist approach is more conservative than Bayesian statistics: the region corresponding to the desired confidence level (frequentist) or likelihood content (Bayesian) γ, has bounds given by $\chi^2(\boldsymbol{\theta}) = Q_k^\gamma$, with $k = 2$ for Bayesian statistics and two-dimensional plots, and $k = f$ for frequentist statistics independently on the number of parameters considered. Since in general and for reasonably good ML values $\chi^2_{\rm min} \lesssim \mathcal{O}(f)$ and $f > 2$, we have that the probability/likelihood content is the same, i.e.

$$\int_{Q_f^\gamma}^{\infty} \mathcal{P}_{\chi_f^2}(u)\mathrm{d}u = \int_{Q_2^\gamma}^{\infty} \mathcal{P}_{\chi_2^2}(u)\mathrm{d}u \tag{5.64}$$

only for $Q_f^\gamma > Q_2^\gamma$. When looking at Bayesian likelihood contours one should thus keep in mind that a point more than, say, 3σ away from the ML point is not necessarily ruled out by data. In order to establish this, one has to look at confidence contours, i.e. ask the frequentist's question. This is pointed out in a penetrating way by Gawiser (2001).

Maximization instead of marginalization

In practical applications, involving up to a dozen parameters, it is an exceptionally demanding task to perform the multidimensional integral of Eq. (5.58). A computationally more feasible alternative which avoids the time consuming integration is to maximize the parameters we are not interested in, $\mathbf{u}$, for each value of the parameters of interest, $\mathbf{t}$, obtaining

$$L(\mathbf{t}) \propto \max_{\mathbf{u}} L(\mathbf{t}, \mathbf{u})\,. \tag{5.65}$$

If the distribution is Gaussian, then the two procedures give the same result: maximizing $L(\mathbf{t}, \mathbf{u})$ corresponds to minimization over $\mathbf{u}$ of the quadratic form $\boldsymbol{\theta}^T \mathbf{C}^{-1} \boldsymbol{\theta}$, with the notations of (5.53) and (5.54). Differentiating with respect to $\mathbf{u}$, we find that the minimum of the quadratic form lies at

$$\mathbf{u} = -\mathbf{B}^{-1}\mathbf{G}^T\mathbf{t}\,, \tag{5.66}$$

and therefore

$$L(\mathbf{t}) \propto \exp -\frac{1}{2}\mathbf{t}^T \left[\mathbf{A} - \mathbf{G}\mathbf{B}^{-1}\mathbf{G}^T\right] \mathbf{t}\,, \tag{5.67}$$

which is the same result we found by marginalizing over $\mathbf{u}$, Eq. (5.60). Numerical investigations have found that maximization tends to underestimate errors when the assumption of a Gaussian distribution is not accurately fulfilled (Efstathiou et al., 1999).

5.1.7 Markov chain Monte Carlo

A big practical limitation to grid based parameter extraction techniques is that the number of CMB spectra needed scales exponentially with the dimensionality of the parameter space

considered. Even with fast parallel computing, the required computational time quickly becomes very large, even for a moderate number of points in each dimension. Interpolation algorithms and other optimization techniques have been employed to circumvent this fundamental limitation, allowing the handling of up to a dozen parameters (Tegmark et al., 2001). Nevertheless, this method shows a lack of flexibility if one wants to add new data-sets or incorporate new parameters or theoretical priors. At the latest with the coming of WMAP data, the days of grid-based parameter extraction seem to be over, since the accuracy of WMAP-like data cannot be exploited with the insufficient resolution and flexibility offered by this technique.

Markov chain Monte Carlo (hereafter MCMC) methods are now becoming the standard tool to determine parameters from CMB data, combine it with large scale structure constraints or investigate the effect of different priors. As advocated e.g. by Christensen et al. (2001), MCMC is a method to generate a sequence of (correlated) samples, called a Markov chain, from the posterior pdf of the parameters given the data, (5.33, page 118). The great advantages are that the computational time scales approximately linearly with the number of dimensions of the parameter space, and that once the chain has properly converged (see below for more details), the marginalized posterior distribution for the parameter(s) of interest can be simply recovered by plotting histograms of the sample list, thus avoiding completely the costly integration. It is easy to adjust the prior information or to include new data-sets into an existing chain without having to recompute it, with a procedure called "importance sampling".

One can think of the MCMC algorithm as an efficient integration technique to evaluate the posterior distribution in Bayes' Theorem, Eq. (5.33, page 118). The Monte Carlo sampling does not rely on the assumption of Gaussian pdf's: indeed, the direct sampling of the posterior permits to reveal features due to its non-Gaussian distribution, and therefore vastly improves on the methods based on chi-square goodness-of-fit described above. Besides those undeniable advantages over the grid method, the popularity of MCMC in the cosmology community has been boosted by the timely public release of the COSMOMC package (Lewis & Bridle, 2002), which integrates the code CAMB for the computation of the CMB power spectra[4] and several useful tools for the generation and interpretation of Markov chains using CMB and other cosmological data-sets. Further details about MCMC methods can be found e.g. in Gilks et al. (1996); MacKay (2003).

The Metropolis-Hastings algorithm (Metropolis et al., 1953; Hastings, 1970) is the core of the sample generation, and produces a Markov chain whose equilibrium distribution is the target probability density, here the posterior $\mathcal{P}(\boldsymbol{\theta}|\mathbf{d})$. The chain is started from a random point in parameter space, $\boldsymbol{\theta}_0$, and a new point $\boldsymbol{\theta}_1$ is proposed with an arbitrarily *proposal density distribution* $q(\boldsymbol{\theta}_n, \boldsymbol{\theta}_{n+1})$. The *transition kernel* $T(\boldsymbol{\theta}_n, \boldsymbol{\theta}_{n+1})$ gives the conditional probability for the chain to move from $\boldsymbol{\theta}_n$ to $\boldsymbol{\theta}_{n+1}$, and it must satisfy the "detailed balance"

$$\mathcal{P}(\boldsymbol{\theta}_{n+1}|\mathbf{d})T(\boldsymbol{\theta}_{n+1}, \boldsymbol{\theta}_n) = \mathcal{P}(\boldsymbol{\theta}_n|\mathbf{d})T(\boldsymbol{\theta}_n, \boldsymbol{\theta}_{n+1}) \tag{5.68}$$

so that the posterior $\mathcal{P}(\boldsymbol{\theta}|\mathbf{d})$ is the stationary distribution of the chain. This is achieved by

[4]Both codes are available at: `http://cosmologist.info`.

defining the transition kernel as

$$T(\boldsymbol{\theta}_n, \boldsymbol{\theta}_{n+1}) \equiv q(\boldsymbol{\theta}_n, \boldsymbol{\theta}_{n+1})\alpha(\boldsymbol{\theta}_n, \boldsymbol{\theta}_{n+1})\,, \tag{5.69}$$

$$\alpha(\boldsymbol{\theta}_n, \boldsymbol{\theta}_{n+1}) \equiv \min\left\{1, \frac{\mathcal{P}(\boldsymbol{\theta}_{n+1}|\mathbf{d})q(\boldsymbol{\theta}_{n+1}, \boldsymbol{\theta}_n)}{\mathcal{P}(\boldsymbol{\theta}_n|\mathbf{d})q(\boldsymbol{\theta}_n, \boldsymbol{\theta}_{n+1})}\right\}\,, \tag{5.70}$$

where $\alpha(\boldsymbol{\theta}_n, \boldsymbol{\theta}_{n+1})$ gives the probability that the new point is accepted. Since $\mathcal{P}(\boldsymbol{\theta}|\mathbf{d}) \propto L(\mathbf{d}|\boldsymbol{\theta})\mathcal{P}(\boldsymbol{\theta})$ and for the usual case of a symmetric proposal density, $q(\boldsymbol{\theta}_n, \boldsymbol{\theta}_{n+1}) = q(\boldsymbol{\theta}_{n+1}, \boldsymbol{\theta}_n)$, the new step is always accepted if it improves on the posterior, otherwise it is accepted with probability $L(\mathbf{d}|\boldsymbol{\theta}_{n+1})\mathcal{P}(\boldsymbol{\theta}_{n+1})/L(\mathbf{d}|\boldsymbol{\theta}_n)\mathcal{P}(\boldsymbol{\theta}_n)$.

The result is a sample list from the target distribution, from which all the statistical quantities of interest can readily be evaluated. The samples are correlated with each other, a fact which does not constitute a problem for the statistical inference on the parameters; however, importance sampling does require uncorrelated samples, which can be obtained from the original chain by suitably "thinning" the chain, i.e. by retaining only one sample every N, with N of the order of a few thousands. Other important practical issues in working with MCMC methods involve:

- **Burn in period:** the initial samples need to be discarded, since the chain is not yet sampling from the equilibrium distribution. The burn in can roughly be assessed by looking at the evolution of the posterior and at the position of the chain in parameter space as a function of the step number. When the chain is started at a random point of the parameter space, the logarithm of the posterior pdf is large (and thus the posterior probability is small), and becomes smaller at every step as the chain approaches the region where the fit to the data is better. Only when the chain has moved in the neighborhood of the ML point the curve of the log posterior as a function of the step number flattens around the best fit value. This is illustrated in the left panel of Fig. 5.1. Another useful diagnostic is the evolution in parameter space of multiple chains, which are started from different points. In a well-behaved situation all of the chains converge after the burn-in period to the same region around the ML point, see the right panel of Fig. 5.1 for an illustration.

- **Convergence:** assessing convergence of the chain essentially means to know when we can stop, having gathered a number of samples large enough to correctly derive the statistical quantities of interest. This is in general a difficult question, see e.g. Cowles & Carlin (1996); Mengersen et al. (1999) and references therein. The COSMOMC package offers several useful diagnostic tools, including the Raftery & Lewis (1996) statistics and the Gelman & Rubin (1992b) criterion.

- **Multiple chains:** there is a debate among experts about the best strategy between having one long chain or rather several shorter ones running in parallel, see e.g. Gelman & Rubin (1992a,b); Raftery & Lewis (1996). Multiple independent chains offer the advantage of being computed in parallel, and can be started in different points of the parameter space to ensure good mixing, i.e. an adequate exploration of the whole parameter space.

- **Starting points:** after the burn in period, the converged chains do not depend on the initial starting points. However, it is convenient to start the chains in the proximity of

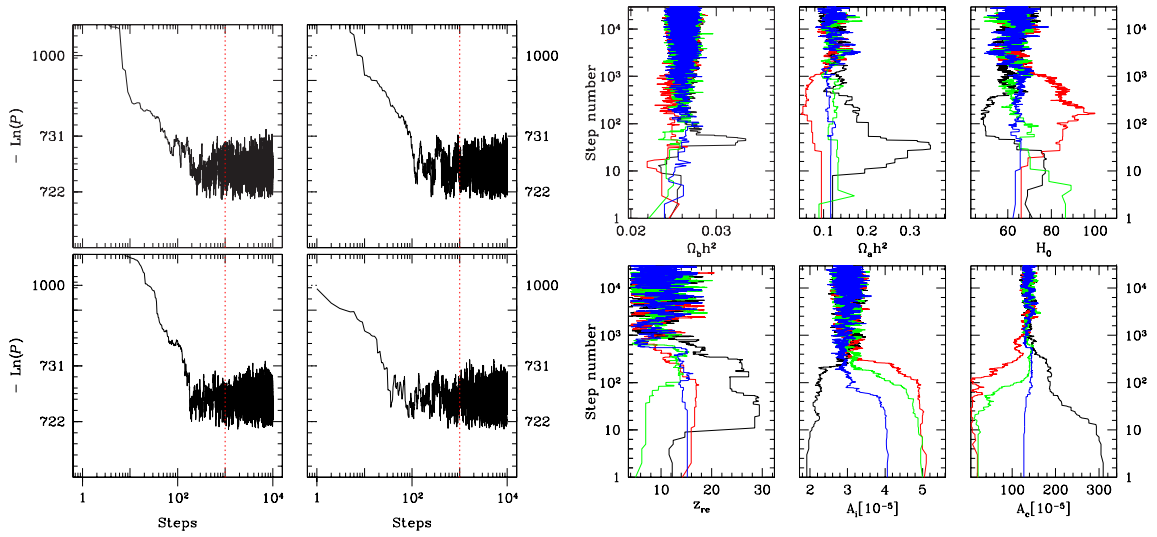

Figure 5.1: Illustration of the burn-in period. Left panel: the logarithm of the (non-normalized) posterior, $-\ln \mathcal{P}(\boldsymbol{\theta}|\mathbf{d})$, as a function of the step number for four Monte Carlo chains. After the burn-in period (dotted, vertical lines), the value flattens and the chains are sampling from the target distribution. Right panel: the four chains (in different colors) are started in different points of a 6-dimensional parameter space and all converge to the same region after the burn-in. The vertical axis gives the number of steps.

the parameter region where the best fit is supposedly located, so that convergence will be quickly achieved, and the sophisticated choice of the starting points proposed by Gelman & Rubin (1992b) is usually not necessary in cosmological applications. Also one has to take into account the fact that the MCMC is a local algorithm, which can be trapped inside local minima far away from the global minimum of the posterior, an issue which is intimately related with the choice of the proposal density. The use of simulated annealing algorithm via the introduction of a finite temperature for the MC can sometimes help in achieving convergence in a weird-shaped parameter space.

- **Proposal density:** the optimal choice of the proposal density is the key parameter for an efficient implementation of the MCMC method. A simple possibility for the proposal density $q(\boldsymbol{\theta}_n, \boldsymbol{\theta}_{n+1})$ is a Gaussian with step size s_i along the parameter direction i, independently on the chain position. Finding the optimal value of s_i is a trade-off between a large step size, which will result in almost all step being rejected and therefore in low efficiency, and a too small value, for which the chain performs a random walk and the tails of the distribution will not be adequately sampled, giving serious underestimate of the likelihood intervals for the parameters. One can also roughly sample the distribution with a short chain, construct from the samples the covariance matrix of the posterior distribution and use this information to construct a new parameter basis approximately aligned with the degeneracy directions (Lewis & Bridle, 2002), which ensures a more efficient exploration. A sampling method which exploits the known degeneracies of the CMB and uses normal parameters as basis has been proposed by

Slosar & Hobson (2003), and it can dramatically enhance the efficiency of the MCMC algorithm, especially for large data-sets as the one expected for the Planck satellite.

5.2 Fisher matrix forecasts

An important issue is to assess quantitatively the expected performance of future CMB experiments in terms of the precision reached in the determination of cosmological parameters: this helps in understanding whether an observed degeneracy is a consequence of the lack of precision in the data, or else it is of fundamental nature and will not be lifted by upcoming or even ideal (i.e. cosmic variance limited) measurements; it also gives estimates of the necessary instrumental characteristics to achieve a certain precision, and on the optimal observing strategies, e.g. full sky coverage versus high resolution mapping of a patch only.

It is possible and indeed necessary at the development stage of a CMB experiment to investigate in detail the above questions by producing mock realizations of the CMB sky and run Monte Carlo simulations of the observations. From the theorist's point of view, however, it is often sufficient and preferable to resort to a simpler alternative, which gives quantitative and accurate results with very small computational requirements: a Fisher matrix analysis (FMA) (Knox, 1995; Kosowsky et al., 1996; Tegmark et al., 1997; Zaldarriaga et al., 1997; Bond et al., 1997; Eisenstein et al., 1998b; Efstathiou & Bond, 1999; Tegmark et al., 2000).

5.2.1 Experimental parameters

As explained in § 5.1.5, if the LF is a multivariate Gaussian then the Fisher information matrix defined in Eq. (5.41) is an estimate of the inverse of the covariance matrix for the parameters under scrutiny. Since any LF can be expanded up to second order in the vicinity of the ML point as in (5.40), the goal is to compute the Fisher matrix for the CMB power spectrum, including the noise of the future experiment, and estimate from it the covariance matrix using the results for Bayesian statistics presented in § 5.1.5.

The estimator (5.15) for the CMB temperature power spectrum (below we generalize the result to include polarization information as well, § 5.2.2) needs to be modified to subtract off the noise contribution and correct for the fact that the measured $a_{\ell m}$'s are a smeared out version of the true ones, resulting from the convolution of the signal with the experimental beam, giving (Knox, 1995; Bond et al., 1997)

$$\hat{C}_\ell \equiv \left(\frac{1}{2\ell+1} \sum_{m=-\ell}^{\ell} |\hat{a}^2_{\ell m}| - w_b^{-1} \right) e^{\ell(\ell+1)/\ell_b^2} \,. \tag{5.71}$$

In the above expression, the two experimental parameters are the *inverse weight per solid angle* w_b, which accounts for the experimental noise, and the *beam width* ℓ_b, which corrects the smoothing due to the Gaussian profile of the beam. These two parameters are written in terms of the fundamental specifications of the experiments, namely the rms pixel noise (or sensitivity per resolution element) σ_b and the angular resolution θ_b (FWHM) expressed in degrees as

$$w_b^{-1} = (\sigma_b \theta_b)^2 \quad \text{and} \quad \ell_b = \sqrt{8 \ln 2}/\theta_b \,. \tag{5.72}$$

In the limit of infinite resolution, $\theta_b \to 0$, and no experimental noise, $\sigma_b \to 0$, we recover the cosmic variance limited estimator (5.15).

As in § 5.1.2, we can now find the pdf for (5.71),

$$\mathcal{P}(\hat{C}_\ell) = \frac{l}{C_\ell + w_b^{-1} e^{\ell(\ell+1)/\ell_b^2}} \mathcal{P}_{\chi_l^2}\left(l \frac{\hat{C}_\ell + w_b^{-1} e^{\ell(\ell+1)/\ell_b^2}}{C_\ell + w_b^{-1} e^{\ell(\ell+1)/\ell_b^2}} \right) , \tag{5.73}$$

recalling $l \equiv 2\ell + 1$ and the chi-square distribution displayed in Eq. (5.17). The correction for the noise and the beam size makes this estimator biased, i.e.

$$\langle \hat{C}_\ell \rangle = C_\ell + w_b^{-1} e^{\ell(\ell+1)/\ell_b^2} , \tag{5.74}$$

which is exactly what we need to compensate for the experimental noise and beam width. From this it follows from (5.23) and (5.38) that the log-normal LF has the form

$$\mathcal{L}(\boldsymbol{\theta}) = \sum_\ell l \left[\ln\left(C_\ell(\boldsymbol{\theta}) + w_b^{-1} e^{\ell(\ell+1)/\ell_b^2} \right) + \frac{\hat{C}_\ell}{C_\ell(\boldsymbol{\theta}) + w_b^{-1} e^{\ell(\ell+1)/\ell_b^2}} \right] \tag{5.75}$$

and we have dropped several normalization factors which do not depend on $\boldsymbol{\theta}$. Using (5.74) we then obtain for the Fisher information matrix defined in (5.41)

$$F_{ij} = \sum_{\ell=\ell_{\min}}^{\ell_{\max}} \frac{1}{(\Delta C_\ell)^2} \frac{\partial C_\ell}{\partial \theta_i} \frac{\partial C_\ell}{\partial \theta_j} \Big|_{\boldsymbol{\theta}^*} , \tag{5.76}$$

where the quantity $(\Delta C_\ell)^2$ is the standard deviation on the estimate of C_ℓ, and takes into account both the cosmic variance and the experimental error,

$$(\Delta C_\ell)^2 = \frac{2}{2\ell + 1} \left(C_\ell + w_b^{-1} e^{\ell(\ell+1)/\ell_b^2} \right)^2 . \tag{5.77}$$

The sum over multipoles runs over the multipole coverage of the experiment, between $\ell_{\min}$ and $\ell_{\max}$.

Thus once the experimental parameters are specified, the computation of the Fisher matrix only requires the knowledge of the derivatives of the power spectrum with respect to the cosmological parameters. The derivatives are determined numerically as double sided derivatives, see § 5.2.3, and this requires the computation of $2p + 1$ spectra only for p parameters, which is a very small computational effort compared with the full numerical exploration of the likelihood surface.

5.2.2 Generalizations

In this section, we develop the necessary general machinery which refines the above results including a more detailed experimental parametrization and polarization information.

Most experiments present several frequency channels, each of them characterized by its own sensitivity $\sigma_c^{T,P}$ and angular resolution $\theta_c^{T,P}$, both for temperature (T) and E-polarization (P). Furthermore, even full-sky experiments only cover a fraction of the sky, since point source subtraction, foreground removal and galactic plane cuts have to be performed on the full-sky maps. This can be approximately taken into account by assigning a "clean" fraction $f_{\rm sky}$ to the experimental coverage. These factors are accounted for by generalizing the expression (5.77) to (Efstathiou & Bond, 1999)

$$(\Delta C_\ell)^2 = \frac{2}{(2\ell + 1) f_{\rm sky}} \left(C_\ell + B_\ell^{-2} \right)^2 , \tag{5.78}$$

where the *inverse noise term* B_ℓ is given by

$$B_\ell^2 \equiv \sum_c w_c e^{-\ell(\ell+1)/\ell_c^2} \tag{5.79}$$

and w_c, ℓ_c are given by (5.72) for each channel c.

In the more general case, we also want to include E polarization and temperature-polarization correlation (C) along with temperature information: then instead of a single derivative we have a vector of three derivatives with the weighting given by the the inverse of the covariance matrix of the spectra, and the Fisher matrix is given by (Zaldarriaga & Seljak, 1997),

$$F_{ij} = \sum_{\ell=\ell_{\rm min}}^{\ell_{\rm max}} \sum_{X,Y} \frac{\partial C_{X\ell}}{\partial\theta_i} {\rm Cov}^{-1}(C_{X\ell}C_{Y\ell}) \frac{\partial C_{Y\ell}}{\partial\theta_j}\Big|_{\boldsymbol{\theta}^*} \tag{5.80}$$

where ${\rm Cov}^{-1}$ is the inverse of the covariance matrix for the spectra evaluated at the ML point $\boldsymbol{\theta}^*$, θ_i are the cosmological parameters we want to estimate and X, Y stands for T (temperature), E (polarization mode), or C (cross-correlation of the power spectra for T and E).

For each ℓ one has to invert the covariance matrix and sum over X and Y. The diagonal terms of the covariance matrix between the different estimators are given by

$$\begin{aligned}
{\rm Cov}(C_{T\ell}^2) &= \frac{2}{(2\ell+1)f_{\rm sky}}(C_{Tl} + B_{T\ell}^{-2})^2 & (5.81)\\
{\rm Cov}(C_{E\ell}^2) &= \frac{2}{(2\ell+1)f_{\rm sky}}(C_{E\ell} + B_{P\ell}^{-2})^2 & (5.82)\\
{\rm Cov}(C_{C\ell}^2) &= \frac{1}{(2\ell+1)f_{\rm sky}}\left[C_{C\ell}^2 + (C_{T\ell} + B_{T\ell}^{-2})(C_{E\ell} + B_{P\ell}^{-2})\right], & (5.83)
\end{aligned}$$

and the off diagonal terms are

$$\begin{aligned}
{\rm Cov}(C_{T\ell}C_{E\ell}) &= \frac{2}{(2\ell+1)f_{\rm sky}}C_{C\ell}^2 & (5.84)\\
{\rm Cov}(C_{T\ell}C_{C\ell}) &= \frac{2}{(2\ell+1)f_{\rm sky}}C_{C\ell}(C_{T\ell} + B_{T\ell}^{-2}) & (5.85)\\
{\rm Cov}(C_{E\ell}C_{C\ell}) &= \frac{2}{(2\ell+1)f_{\rm sky}}C_{C\ell}(C_{E\ell} + B_{P\ell}^{-2}), & (5.86)
\end{aligned}$$

where $B_{T\ell}^{-2} = B_\ell^{-2}$ given in Eq. (5.79) and $B_{P\ell}^2$ is obtained using a similar expression but with the experimental specifications for the polarization channels.

5.2.3 Accuracy issues

The accuracy of the Fisher matrix predictions for the errors depends on a number of issues:

- The FMA *assumes* that the true values of the parameters are in the vicinity of the ML point $\boldsymbol{\theta}^*$. The validity of the results therefore depends on this assumption, as well as on the assumption that the $a_{\ell m}$'s are independent Gaussian random variables.

- This is a local method based on a quadratic expansion of the LF. Only if the FMA predicted errors are small enough, the method is self-consistent and we can expect

the FMA prediction to correctly reproduce the exact behavior, and in particular the correlations between parameters, thus revealing the degeneracy directions. The expansion up to second order is *exact* if the dependence of the C_ℓ on the parameters is linear, therefore great importance is attached to the choice of the parameter set with respect to the FMA is performed. As shown in Kosowsky et al. (2002), employing the normal parameters set discussed in § 4.2 as a base, the accuracy of the FMA predictions is greatly enhanced. This is because the spectra are almost linear in the normal parameters in the vicinity of the best fit.

- Special care must be taken when computing the derivatives of the power spectrum with respect to the cosmological parameters. This differentiation strongly amplifies any numerical errors in the spectra, leading to larger derivatives, which would artificially break degeneracies among parameters. Double–sided derivatives reduce the truncation error from second order to third order terms, but the correct choice of the step size is a trade-off between truncation error and numerical inaccuracy dominated cases (Press et al., 1992).

5.3 CMB observations: a brief historical account

The experimental status of CMB observations has made giant leaps over the last ten years, thanks to spectacular advancements in detector technology. As demonstrated in Chapter 6, CMB data nowadays provide stringent tests which severely constrain cosmological model building, and call for more refined theoretical and computational approaches which take into account subtle physical effects which were so far ignored or thought to be irrelevant. Here we provide a personal selection of a few milestones of this development, in order to put the current and future experimental achievements into a wider perspective.

The first detection of temperature anisotropy came in 1992 with the Differential Microwave Radiometer (DMR) aboard the COBE satellite after one year of observations on angular scales larger than 7° (Smoot et al., 1992; Wright et al., 1992) or multipoles $\lesssim 20$. The key results of the full four year DMR observations are summarized in Bennett et al. (1996, see references therein): the quadrupole amplitude was measured for the first time, the spectral tilt of the large scale spectrum was found to be compatible with an Harrison-Zel'dovich spectrum and no evidence of non-Gaussianity of the fluctuations was discovered in the data. The FIRAS instrument was devoted to the study of the CMB spectrum (Fixsen et al., 1996), and obtained a precision measurement of its temperature ($T = 2.728 \pm 0.002$ K), while constraining deviations from a perfect black body spectrum to be less than about one part in 10^5 with 95% confidence.

The Saskatoon and Toco data provided the first hint for the presence of the first adiabatic peak (Netterfield et al., 1997; Miller et al., 1999; Knox & Page, 2000), but at the turning of the millennium several groups independently reported measurements of the temperature anisotropy with a resolution of a few arcminutes, sufficient to unambiguously reveal the first peak and start exploring the subsequent ones: BOOMERanG (de Bernardis et al., 2002; Netterfield et al., 2002) and Maxima (Hanany et al., 2000; Lee et al., 2001), both balloon-borne bolometric experiments, mapped the multipole region $80 \lesssim \ell \lesssim 1000$; the CBI (Padin et al., 2001) and DASI (Halverson et al., 2002) ground based interferometers covered a sim-

ilar multipole range but with a completely different technology, which had the advantage of being free from the calibration uncertainty of bolometric receivers. The Archeops experiment (Benoit et al., 2003a), conceived as a balloon-borne precursor of the HFI bolometric instrument for the Planck satellite, observed a larger portion of the sky, and thus provided an estimation of the temperature power spectrum which for the first time encompassed the first peak region and also partially overlapped with the COBE measurement, in the range $15 \leq \ell \leq 350$. Given the experimental calibration uncertainty of the bolometers, which is about $10 - 20\%$, this permits to test the relative calibration between COBE and the other experiments with data in the $\ell \gtrsim 50$ region, and perform a comparison of the height of the first peak with respect to the large scale plateau. All of this data generally agrees well on the position and shape of the first peak, but their resolution is insufficient to permit the reconstruction of the subsequent ones with high confidence (de Bernardis et al., 2002; Durrer et al., 2003b).

From the point of view of parameter extraction, each of the above data sets by its own as well as their combination leads to a broad agreement of an approximately flat $\Omega_{\rm tot} \sim 1$ universe with scale invariant spectral index $n_s \sim 1$, with the 1σ likelihood intervals being of the order of 10% and somewhat depending on the compilation of data and on the prior assumed (Stompor et al., 2001; Lange et al., 2001; Pryke et al., 2002; Netterfield et al., 2002). The estimation of the baryon density proved to be more controversial, because of discrepancies and a lack of resolution at the level of the second and third peak: in particular, the BOOMERanG 1998 and MAXIMA data seem to favor a baryon content about 50% larger than predicted by BBN, around $\Omega_b h^2 \sim 0.03$ (Tegmark & Zaldarriaga, 2000; Lange et al., 2001; Stompor et al., 2001), a discrepancy which disappears with the improved beam reconstruction of the BOOMERanG 2000 observations (Netterfield et al., 2002). Inclusion of supernovæ data or the Hubble Space Telescope prior for the Hubble constant, together with the flatness determination, points toward a universe dominated by a cosmological constant.

Before the WMAP satellite delivered its results, ground based instruments pressed on and opened up two new observational directions: very small scale observations ($4' - 5'$)

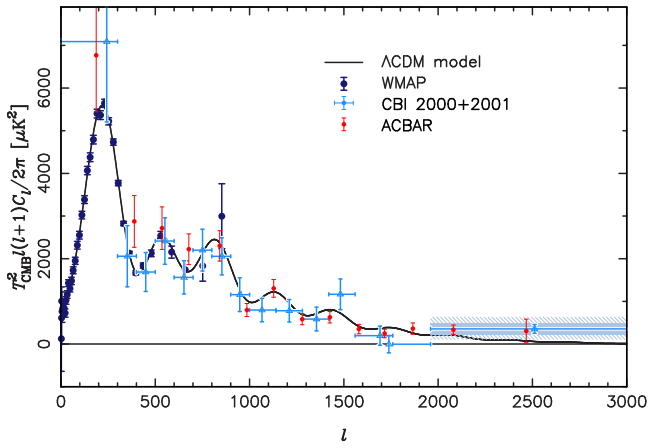

Figure 5.2: The small scale temperature angular power spectrum observed by CBI "mosaic" during two years and by ACBAR. The shaded region shows the excess power at small scale, compatible with the SZ effect. Reprinted from Readhead et al. (2004).

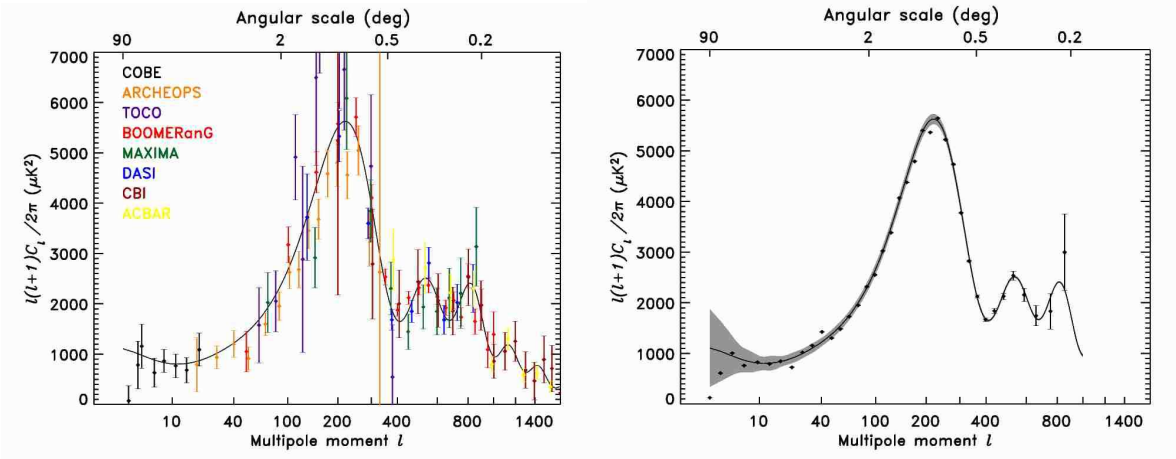

Figure 5.3: The spectacular increase of the accuracy of CMB observations: in the left panel, a compilation of pre-WMAP temperature power spectrum measurements obtained between 1996 (COBE) and 2003 (CBI) is compared with the WMAP first year data in the right panel, released in February, 2003. The error-bars give the 1σ uncertainty due to the measurement errors, while the shaded region represent the cosmic variance limit. Both figures reprinted from Hinshaw et al. (2003a).

and E-polarization detection. The CBI interferometer, in two different configurations called "mosaic" and "deep field", obtained measurements of the temperature power spectrum up to $\ell = 3500$ (Sievers et al., 2003; Mason et al., 2003), and it was argued that the excess power observed at high multipoles could be due to the SZ effect, from which a precise determination of σ_8 could possibly be obtained (Bond et al., 2002). The ACBAR experiment, a bolometric instrument installed at the South Pole, found small scale power consistent with the results of CBI, without however being able to place tighter constraints on its origin (Goldstein et al., 2003; Kuo et al., 2004). More recently, the results of two years of observations with the CBI "mosaic" configuration, give smaller errors in the $\ell \sim 2000$ region, due to the longer integration time and to an improved absolute calibration derived from the WMAP data, see Fig. 5.2. Beside revealing effects due to secondary anisotropies as the SZ effect, the small scale measurements are helpful in better constraining $n_{\rm s}$, $\tau_{\rm re}$ and possible features in the power spectrum (like a "running", i.e. a scale dependence of $n_{\rm s}$) because of the larger lever arm they offer when combined with WMAP and large scale structure data (Readhead et al., 2004).

The DASI interferometer reported in the second half of 2002 the first detection of E-polarization, which was observed on degree angular scales with almost 5σ confidence (Kovac et al., 2002), thereby opening the epoch of polarization measurements.

The first year WMAP data, unveiled in February 2003 (Bennett et al., 2003; Hinshaw et al., 2003a), essentially confirmed the picture which had emerged from pre-WMAP observations, see Fig. 5.3: the height of the first peak was corrected by about 10%, showing more power than in the previous data, while the large scale spectrum confirmed the DMR results. The second peak is now accurately outlined, while the full four years data should allow to obtain good resolution up to $\ell \sim 1000$ in temperature. The low power of the quadrupole remains troublesome, since it is still not clear whether it is pointing to new physics or just a

consequence of systematical errors. The observation of the temperature-polarization correlation up to $\ell \sim 500$ (Kogut et al., 2003) has proved very useful in order to better constrain parameters. The exquisite quality of the power spectra has tightened the 1σ likelihood intervals to a few percent for most cosmological parameters (Spergel et al., 2003), and the central value has remained in the region preferred by earlier data, with two interesting exceptions: the TE data favor a much larger reionization optical depth than previously thought, and there seems to be a slight preference for a "running" (i.e. scale dependent) spectral index (Peiris et al., 2003).

A complete overview of the evolution of data and of the cosmological parameters derived from it can be found in the review by Bond et al. (2003).

Chapter 6

Beyond standard parameters

This chapter is devoted to the investigation of three scenarios involving non-standard cosmological parameters, and focuses on the ability of constraining them using present and future CMB observations: the existence of extra relativistic particles (§ 6.1); the determination of the primordial helium mass fraction (§ 6.2); and possible time variations of the fine structure constant (§ 6.3).

Until recently, the effects induced by these parameters on the CMB where considered too small to be observable, or else irrelevant; however, the era of precision cosmology that we are entering requires on one hand that we check the consequences of our assumptions on the standard results for other parameters (as in the case of the neutrino families and the helium fraction); on the other hand, it allows us to put under close scrutiny very subtle effects which could previously be safely neglected because of the less accuracy of the data sets.

6.1 Extra relativistic particles

This section is based on the work published in Bowen et al. (2002), which was carried out for the most part during my stay in Oxford. We investigate one possible modification to the standard scenario, namely variations in the parameter $\omega_{\rm rel} = \Omega_{\rm rel} h^2$ which describes the energy density of relativistic particles. The original work has been performed in 2001, and therefore the results presented here of the pre-WMAP data analysis are nowadays somewhat outdated. However, the focus is on the degeneracies involving $\omega_{\rm rel}$ and as such the conclusions drawn are still valid. Furthermore, the subsequent analysis by several groups of the actual WMAP data permits a comparison between the forecasts obtained with the Fisher matrix technique in 2001 and the real case, showing a very satisfactory agreement and validating the method used.

After offering the motivations for our study in § 6.1.1, we review various physical mechanisms that can lead to a change in $\omega_{\rm rel}$ with respect to the standard value in § 6.1.2. In § 6.1.3, we illustrate how the CMB angular power spectrum depends on this parameter and identify possible degeneracies with other parameters, then present in § 6.1.4 a likelihood analysis from pre-WMAP CMB data and show which of the constraints on the various parameters are affected by variations in $\omega_{\rm rel}$. Section 6.1.5 forecasts the precision in the estimation of cosmological parameters for the space missions WMAP and Planck, and then compares the predictions with actual data analysis performed on the first year WMAP data.

6.1.1 Motivation

CMB data analysis taking into account variations in the density of relativistic particles has been previously undertaken by many authors (Hannestad, 2000; Esposito et al., 2001; Kneller et al., 2001; Hannestad, 2001; Hansen et al., 2002; Zentner & Walker, 2002), giving rather crude upper bounds, which are significantly improved only by including priors on the age of the universe or by including supernovae (SN) or large scale structure (LSS) data. It is worth emphasizing that there is little difference in the bounds on $N_{\rm eff}$, the effective number of relativistic species, obtained from old and recent CMB data because of the degeneracy described in detail below. We focus here on the effects that the inclusion of this parameter, $\omega_{\rm rel}$, has on the constraints of the remaining parameters in the context of purely adiabatic models.

As shown below – and as observed previously, see e.g. Hu et al. (1999) – there is a strong degeneracy between $\omega_{\rm rel}$ and the physical density of non-relativistic matter, $\omega_m \equiv \Omega_m h^2$. This is important, because an accurate determination of ω_m from CMB observations (and of Ω_m by including the Hubble Space Telescope result $h = 0.72 \pm 0.08$) can be useful for a large number of reasons. First of all, determining the cold dark matter content, $\omega_{\rm cdm} = \omega_m - \omega_b$ can shed new light on the nature of dark matter. The thermally averaged product of cross-section and thermal velocity of the dark matter candidate is related to ω_m, and this relation can be used to analyze the implications for the mass spectra in versions of the Supersymmetric Standard Model, see e.g. Barger & Kao (2001); Djouadi et al. (2001); Ellis et al. (2001). The value of Ω_m can be determined in an independent way from the mass-to-light ratios of clusters, and the present value is $0.1 < \Omega_m < 0.2$ (Carlberg et al., 1997; Bahcall et al., 2000). Furthermore, a precise measurement of Ω_m will be a key input for determining the redshift evolution of the equation of state parameter $w(z)$ and thus discriminating between different quintessential scenarios, see e.g. Weller & Albrecht (2002).

6.1.2 Effective number of relativistic species

The energy density of relativistic particles can conveniently be parameterized via the effective number of relativistic species, $N_{\rm eff}$: in the standard model $\omega_{\rm rel}$ includes photons and neutrinos, and it can be expressed as

$$\omega_{\rm rel} = \omega_\gamma + N_{\rm eff} \cdot \omega_\nu \tag{6.1}$$

where ω_γ is the energy density in photons and ω_ν is the energy density in one active neutrino family. In geometrical units, where $G = \hbar = c = 1$, one has $\omega_x = 4\pi^3/45 \cdot g_x T_x^4$, where g_x and T_x are the relativistic degrees of freedom and the temperature of species $x = \gamma, \nu$, respectively. Measuring $\omega_{\rm rel}$ thus gives a direct observation of the effective number of neutrinos, $N_{\rm eff}$. Naturally there are only three active neutrinos, and $N_{\rm eff}$ is simply a convenient parametrization for the extra possible relativistic degrees of freedom

$$N_{\rm eff} = 3 + \Delta N\,. \tag{6.2}$$

Thus $\omega_{\rm rel}$ includes energy density from all the relativistic particles: photons, neutrinos, and additional hypothetical relativistic particles such as a light majoron or a sterile neutrino. Such hypothetical relativistic particles are strongly constrained from standard Big-Bang

nucleosynthesis (BBN), where the allowed extra relativistic degrees of freedom typically are expressed through the effective number of neutrinos, $N_{\rm eff} = 3 + \Delta N_{\rm BBN}$. BBN bounds are typically about $\Delta N_{\rm BBN} < 0.2 - 1.0$ (Burles et al., 1999; Lisi et al., 1999).

One should, however, be careful when comparing the effective number of neutrino degrees of freedom at the time of BBN (neutrino decoupling) and at the formation of the CMBR (photon decoupling). This is because the energy density in relativistic species may change from the time of BBN ($T \sim$ MeV) to the time of last rescattering ($T \sim$ eV), as explained in Hansen et al. (2002). For instance, if one of the active neutrinos has a mass in the range eV $< m <$ MeV and decays into sterile particles such as other neutrinos, majorons etc. with lifetime t(BBN) $< \tau < t$(CMBR), then the effective number of neutrinos at CMBR would be substantially different from the number at BBN (White et al., 1995). Such massive active neutrinos, however, do not look very natural any longer in view of the recent experimental results on neutrino oscillations (Fogli et al., 2001; Gonzalez-Garcia et al., 2001), showing that all active neutrinos are likely to have masses smaller than 0.1 eV. One could instead consider sterile neutrinos mixed with active ones which could be produced in the early universe by scattering, and subsequently decay. The mixing angle must then be large enough to thermalize the sterile neutrinos, and this can be expressed through the sterile to active neutrino number density ratio $n_s/n_\nu \approx 4{\cdot}10^4 \sin^2 2\theta\ (m/\mathrm{keV})(10.75/g^*)^{3/2}$ (Dolgov & Hansen, 2002), where θ is the mixing angle, and g^* counts the relativistic degrees of freedom, such that $n_s/n_\nu = 1$ or $\Delta g^* = 7/8$ increases $N_{\rm eff}$ by one unit. With n_s/n_ν of order unity we use the decay time, $\tau \approx 10^{20}(\mathrm{keV}/m)^5/\sin^2 2\theta$ sec, and one finds, $\tau \approx 10^{17}(\mathrm{keV}/m)^4$ yr, which is much longer than the age of the universe for $m \sim$ keV, so they would certainly not have decayed at t(CMBR). A sterile neutrino with a mass of a few MeV would seem to have the right decay time, $\tau \sim 10^5$ yr, but this is excluded by standard BBN considerations (Kolb et al., 1991; Dolgov et al., 1998). More inventive models with particles decaying during last rescattering cannot simply be treated with an $N_{\rm CMB}$ that is constant in time, see e.g. Kaplinghat et al. (1999), and we will not discuss such possibilities further here.

Even though the simplest models predict that the relativistic degrees of freedom are the same at BBN and CMB times, one could construct models such as quintessence (Albrecht & Skordis, 2000; Skordis & Albrecht, 2002) which effectively could change ΔN between BBN and CMB (Bean et al., 2001). Naturally ΔN can be both positive and negative. For BBN, ΔN can be negative if the electron neutrinos have a non-zero chemical potential (Kang & Steigman, 1992; Kneller et al., 2001), or more generally with a non-equilibrium electron neutrino distribution function (Hansen & Villante, 2000). To give an explicit (but highly exotic) example of a different number of relativistic degrees of freedom between BBN and CMB, one could consider the following scenario. Imagine another two sterile neutrinos, one of which is essentially massless and has a mixing angle with any of the active neutrinos just big enough to bring it into equilibrium in the early universe, and one with a mass of $m_{\nu_s} = 3$ MeV and decay time $\tau_{\nu_s} = 0.1$ sec, in the decay channel $\nu_s \to \nu_e + \phi$, with ϕ a light scalar. The resulting non-equilibrium electron neutrinos happen to exactly cancel the effect of the massless sterile state, and hence we have $\Delta N_{\rm BBN} = 0$. However, for CMB the picture is much simpler, and we have just the stable sterile state and the majoron, hence $\Delta N_{\rm CMB} = 1.57$. For CMB, one can imagine a negative ΔN from decaying particles, where the decay products are photons or electron/positrons which essentially increases the photon temperature relative

to the neutrino temperature (Kaplinghat & Turner, 2001). Such a scenario also naturally dilutes the baryon density, and the agreement on ω_b from BBN and CMB gives a bound on how negative $\Delta N_{\rm CMB}$ can be. Considering all these possibilities, we will therefore not make the usual assumption, $\Delta N_{\rm BBN} = \Delta N_{\rm CMB}$, but instead consider $\Delta N_{\rm CMB}$ as a completely free parameter in the following analysis.

The standard model value for $N_{\rm eff}$ with three active neutrinos is 3.044. This small correction arises from the combination of two effects arising around the temperature $T \sim$ MeV. These effects are the finite temperature QED correction to the energy density of the electromagnetic plasma (Heckler, 1994), which gives $\Delta N = 0.01$ (Lopez & Turner, 1999; Lopez et al., 1999). If there are more relativistic species than active neutrinos, then this effect will be correspondingly higher (Steigman, 2001). The other effect comes from neutrinos sharing in the energy density of the annihilating electrons (Dicus et al., 1982), which gives $\Delta N = 0.034$ (Dolgov et al., 1997; Esposito et al., 2000c). Thus one finds $N_{\rm eff} = 3.044$. An accurate analysis which takes into account both of this effects simultaneously has been performed by Mangano et al. (2002) and the result indicates that the combined effect is slightly smaller, $N_{\rm eff} = 3.0395$.

6.1.3 CMB theory and degeneracies

As explained in detail in Chapter 4, the structure of the C_ℓ spectrum depends on a restricted combination of cosmological parameters, which are physically probed by the CMB; simplifying somewhat the normal parameters set introduced in § 4.2, we focus here on the four cosmological parameters

$$\omega_b \ , \ \omega_m \ , \ \omega_{\rm rel} \text{ and } \mathcal{R}^{\rm shift} \ , \tag{6.3}$$

the physical baryonic density $\omega_b \equiv \Omega_b h^2$, the energy density in matter $\omega_m \equiv (\Omega_{\rm cdm} + \Omega_b)h^2$, the energy density in radiation $\omega_{\rm rel}$ and the shift parameter $\mathcal{R}^{\rm shift} \equiv \ell_{\rm ref}/\ell$, which gives the position of the acoustic peaks with respect to a flat, $\Omega_\Lambda = 0$ reference model, see Eq. (4.22, page 94). In previous analysis (Efstathiou & Bond, 1999; Melchiorri & Griffiths, 2001), the parameter $\omega_{\rm rel}$ has been kept fixed to the standard value, while here we will allow it to vary.

It is therefore convenient to write

$$\omega_{\rm rel} = 4.13 \cdot 10^{-5}(1 + 0.135 \cdot \Delta N_{\rm CMB}) \tag{6.4}$$

(taking $T_{\rm CMB} = 2.726$ K), where $\Delta N_{\rm CMB}$ is the excess number of relativistic species with respect to the standard model, $N_{\rm eff} = 3 + \Delta N_{\rm CMB}$, and we drop the subscript CMB from now on. The shift parameter $\mathcal{R}^{\rm shift}$ depends on $\Omega_m \equiv \Omega_{\rm cdm} + \Omega_b$, on the curvature $\Omega_\kappa \equiv 1 - \Omega_\Lambda - \Omega_m - \Omega_{\rm rel}$, and on $\Omega_{\rm rel} = \omega_{\rm rel}/h^2$ through

$$\mathcal{R}^{\rm shift} = \left(1 - \frac{1}{\sqrt{1+z_{\rm dec}}}\right) \frac{\sqrt{|\Omega_\kappa|}}{\Omega_m} \frac{2}{\chi(\Delta\tau)} \left[\sqrt{\Omega_{\rm rel} + \frac{\Omega_m}{1+z_{\rm dec}}} - \sqrt{\Omega_{\rm rel}}\right] , \tag{6.5}$$

where $z_{\rm dec}$ is a function of the physical baryon density and $\chi(\Delta\tau)$ is given in Eq. (1.33, page 27). Eq. (6.5) generalizes the expression for $\mathcal{R}^{\rm shift}$ given in (4.28, page 95) to the case of non-constant $\Omega_{\rm rel}$.

By fixing the four parameters given in (6.3), or equivalently the set ω_b, the redshift of equality $z_{\rm eq} \equiv \omega_m/\omega_{\rm rel}$, ΔN and $\mathcal{R}^{\rm shift}$, one obtains a perfect degeneracy for the CMB

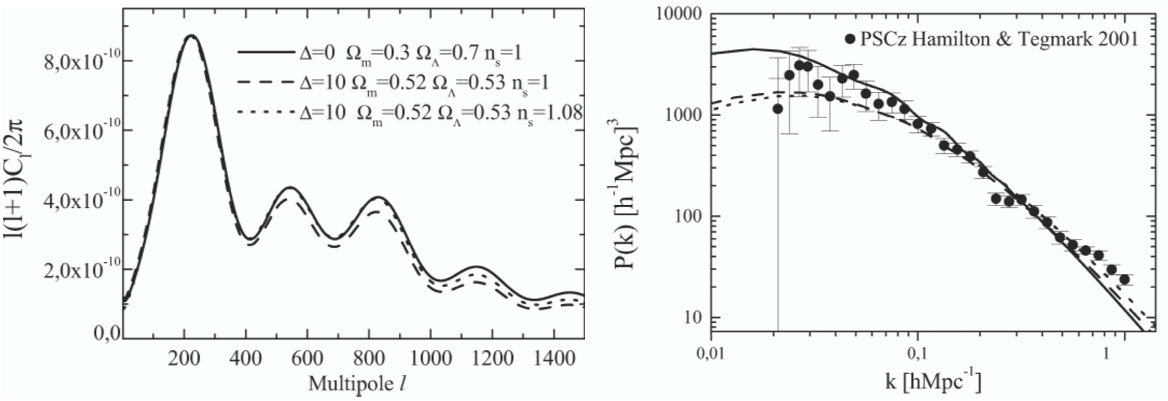

Figure 6.1: Left panel: CMB degeneracies between cosmological models. Keeping $z_{\rm eq}, \omega_b$ and $\mathcal{R}$ fixed while varying ΔN produces nearly degenerate power spectra. The reference model (solid line) has $\Delta N = 0$, $\Omega_{\rm tot} = 1.00$, $n_s = 1.00$; the nearly degenerate model (dotted) has $\Delta N = 10$, $\Omega_{\rm tot} = 1.05$, $n_s = 1.00$. The curves are normalized to the first peak. The position of the peaks is perfectly matched, only the relative height between the first and the other acoustic peaks is somewhat different in this extreme example, due to the early ISW effect. The degeneracy can be further improved, at least up to the third peak, by raising the spectral index to $n_s = 1.08$ (dashed). Right panel: the matter power spectra of the models plotted in the top panel together with the observed decorrelated power spectrum from the PSCz survey (Hamilton & Tegmark, 2002). The geometrical degeneracy is now lifted.

anisotropy power spectra on degree and sub-degree angular scales. On larger angular scales, the degeneracy is broken by the late ISW effect because of the different curvature and cosmological constant content of the models. From the practical point of view, however, it is still very difficult to break the degeneracy, since measurements are limited by cosmic variance on those scales, and because of the possible contribution of gravitational waves.

Allowing ΔN to vary, but keeping constant the other three parameters ω_b, $z_{\rm eq}$, and $\mathcal{R}^{\rm shift}$, we obtain nearly degenerate power spectra which we plot in Fig. 6.1, normalized to the first acoustic peak. The degeneracy in the acoustic peaks region is now slightly spoiled by the variation of the ratio $\Omega_\gamma/\Omega_{\rm rel}$: the different radiation content at decoupling induces a larger (for $\Delta N > 0$) early ISW effect, which boosts the height of the first peak with respect to the other acoustic peaks. Nevertheless, it is still impossible to distinguish between the different models with present (pre-WMAP) CMB measurements and without external priors. Furthermore, a slight change in the scalar spectral index, n_s, can reproduce a perfect degeneracy up to the third peak.

The main result is that, even with a measurement of the first three peaks in the angular spectrum, it is impossible to put bounds on $\omega_{\rm rel}$ alone, even when fixing other parameters such as ω_b. Furthermore, since the degeneracy is mainly in $z_{\rm eq}$, the constraints on ω_m from CMB are also affected, see § 6.1.4.

In Fig. 6.2 we plot the shift parameter $\mathcal{R}^{\rm shift}$ as a function of ΔN, while fixing $\Omega_m = 0.3$ and $\Omega_\Lambda = 0.7$. Increasing ΔN moves the peaks to smaller angular scales, even though the dependence of the shift parameter on ΔN is rather mild. In order to compensate this effect, one has to change the curvature by increasing Ω_m and Ω_Λ. We therefore conclude that the

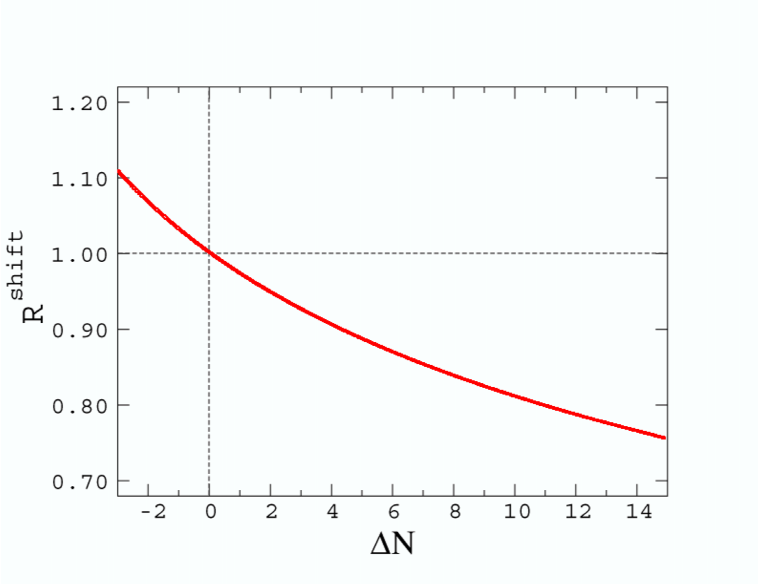

Figure 6.2: The shift parameter $\mathcal{R}^{\rm shift}$ as a function of ΔN with $\Omega_\Lambda = 0.7$ and $\Omega_m = 0.3$. The position of the peaks is only weakly affected by ΔN.

present bounds on the curvature of the universe are weakly affected by ΔN. Nevertheless, when a positive (negative) ΔN is included in the analysis, the preferred models are shifted toward closed (open) universes.

6.1.4 Pre-WMAP constraints from CMB and other data-sets

In this section, we compare pre-WMAP CMB observations with a set of models with cosmological parameters sampled as follows: $0.1 < \Omega_m < 1.0$, $0.1 < \Omega_{\rm rel}/\Omega_{\rm rel}(\Delta N = 0) < 3$, $0.015 < \Omega_b < 0.2$; $0 < \Omega_\Lambda < 1.0$ and $0.40 < h < 0.95$. We vary the spectral index of the primordial density perturbations within the range $n_s = 0.50, ..., 1.50$ and we re-scale the fluctuation amplitude by a pre-factor C_{10}, in units of $C_{10}^{\rm COBE}$. We also restrict our analysis to purely adiabatic, *flat* models ($\Omega_{\rm tot} = 1$) and we add an external Gaussian prior on the Hubble parameter $h = 0.65 \pm 0.2$.

Constraints from CMB only

The theoretical models are computed using the publicly available CMBFAST program (Seljak & Zaldarriaga, 1996) and are compared with the BOOMERanG-98, DASI and MAXIMA-1 data. The power spectra from these experiments were estimated in 19, 9 and 13 bins respectively, spanning the range $25 \leq \ell \leq 1100$. We approximate the experimental signal C_B^{ex} inside the bin to be a Gaussian variable, and we compute the corresponding theoretical value C_B^{th} by convolving the spectra computed by CMBFAST with the respective window functions. When the window functions are not available, as in the case of Boomerang-98, we use top-hat window functions. The likelihood for a given cosmological model is then given by

$$\mathcal{L} = (C_B^{th} - C_B^{ex}) M_{BB'} (C_{B'}^{th} - C_{B'}^{ex}) \tag{6.6}$$

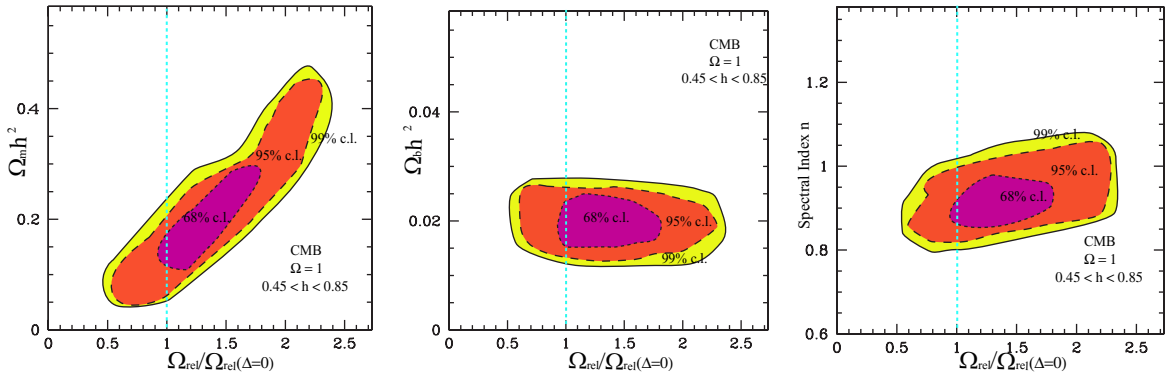

Figure 6.3: Two-dimensional likelihood plots from analysis of CMB data.

where C_B^{th} (C_B^{ex}) is the theoretical (experimental) band power and $M_{BB'}$ is the Gaussian curvature of the likelihood matrix at the peak. This expression is a generalization of Eq. (5.30, page 116) for the case of correlated experimental points. We consider 10%, 4% and 4% Gaussian distributed calibration errors (in μK) for the BOOMERanG-98, DASI and MAXIMA-1 experiments respectively. We also include the COBE data using Lloyd Knox's RADPack package (RADPack Website, 2001).

In order to show the effect of the inclusion of $\omega_{\rm rel}$ on the estimation of the other parameters, we plot likelihood contours in the $\omega_{\rm rel} - \omega_m$, $\omega_{\rm rel} - \omega_b$, $\omega_{\rm rel} - n_s$ planes. Proceeding as in Melchiorri et al. (2000), we calculate a likelihood contour in those planes by maximizing the other parameters as explained in § 5.1.5. In Fig. 6.3 we plot the likelihood contours for $\omega_{\rm rel}$ vs ω_m, ω_b and n_s. As can be seen, $\omega_{\rm rel}$ is very weakly constrained to be in the range $1 \leq \omega_{\rm rel}/\omega_{\rm rel}(\Delta N = 0) \leq 1.9$ at 1σ l.c. in all plots[1]. The degeneracy between $\omega_{\rm rel}$ and ω_m is evident in the left panel of Fig. 6.3. Increasing $\omega_{\rm rel}$ shifts the epoch of matter-radiation equality and this can be compensated only by a corresponding increase in ω_m. It is interesting to note that even if we are restricting our analysis to flat models, the degeneracy is still present and that the bounds on ω_m are strongly affected. We find $\omega_m = 0.2 \pm 0.1$, to be compared with $\omega_m = 0.13 \pm 0.04$ when ΔN is kept to zero. It is important to realize that these bounds on $\omega_{\rm rel}$ appear because of our prior on h and because we consider flat models. When one allows h and Ω_m to be free parameters, then the degeneracy is almost complete and there are no bounds on $\omega_{\rm rel}$.

In the central and right panel of Fig. 6.3 we plot the likelihood contours for ω_b and n_s. As we can see, these parameters are not strongly affected by the inclusion of $\omega_{\rm rel}$. The bound on ω_b, in particular, is completely unaffected by $\omega_{\rm rel}$. There is however, a small correlation between $\omega_{\rm rel}$ and n_s: the boost of the first peak induced by the ISW effect can be compensated (at least up to the third peak) by a small change in n_s (right panel).

Since the degeneracy is mainly in $z_{\rm eq}$, it is useful to estimate the constraints we can put on this variable. In Fig. 6.4 we plot the likelihood curve on $z_{\rm eq}$ alone obtained by maximizing over all other parameters. By integration of this probability distribution function we obtain

$$z_{\rm eq} = 3100^{+600}_{-400} \quad \text{at } 68\% \text{ l.c.} \tag{6.7}$$

[1]Here as in the following, the abbreviation "l.c." stands for "likelihood content", in the Bayesian sense explained in § 5.1.5.

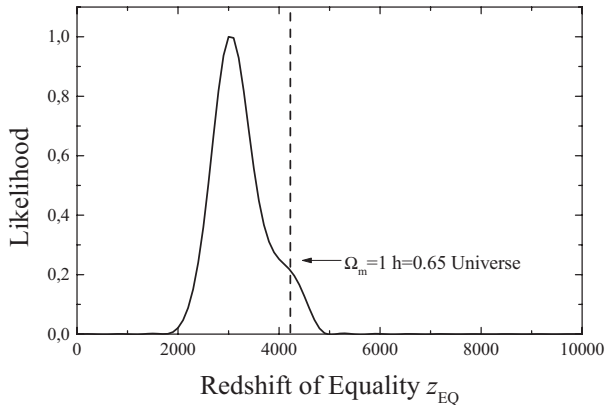

Figure 6.4: Likelihood probability distribution function for the redshift of equality.

Adding other data-sets

It is interesting to investigate how well the constraints from CMB-independent data-sets can break the degeneracy between $\omega_{\rm rel}$ and ω_m. The supernovae luminosity distance is very weakly dependent on $\omega_{\rm rel}$ – see however Zentner & Walker (2002) – and the bounds obtained on Ω_m can be used to break the CMB degeneracy. Including the SN-Ia constraints on the $\Omega_m - \Omega_\Lambda$ plane, $0.8\Omega_m - 0.6\Omega_\Lambda = -0.2 \pm 0.1$ (Perlmutter et al., 1999), we find

$$\omega_{\rm rel}/\omega_{\rm rel}(\Delta N = 0) = 1.12^{0.35}_{-0.42} \quad \text{at } 2\sigma\% \text{ l.c.} \tag{6.8}$$

It is also worth including constraints from galaxy clustering and local cluster abundances. The degeneracy between ω_m and $\omega_{\rm rel}$ in the CMB cannot be broken trivially by inclusion of large-scale structure (LSS) data, because a similar degeneracy affects the LSS data as well (Hu et al., 1999). However, the geometrical degeneracy is lifted in the matter power spectrum, and accurate measurements of galaxy clustering at very large scales can distinguish between various models. This is exemplified in the right panel of Fig. 6.2, where we plot three matter power spectra with the same cosmological parameters as in the top panel, together with the decorrelated matter power spectrum obtained from the PSCz survey.

The shape of the matter power spectrum in the linear regime for galaxy clustering can be characterized by the shape parameter

$$\Gamma \sim \frac{\Omega_m h}{\sqrt{1 + 0.135\Delta N}} e^{-(\Omega_b(1+\sqrt{2h}/\Omega_m) - 0.06)} . \tag{6.9}$$

From the observed data one has roughly (Bond & Jaffe, 1999) $0.15 \leq \Gamma + (n_s - 1)/2 \leq 0.3$. The inclusion of this (conservative) value on Γ gives

$$\omega_{\rm rel}/\omega_{\rm rel}(\Delta N = 0) = 1.40^{0.49}_{-0.56} \quad \text{at } 2\sigma\% \text{ l.c.} \tag{6.10}$$

a bound which is less less restrictive than the one obtained using the SN-Ia prior.

A better constraint can be obtained by including a prior on the variance of matter perturbations over a sphere of size $8h^{-1}$ Mpc, derived from cluster abundance observations. Comparing with $\sigma_8 = (0.55 \pm 0.05)\Omega_m^{-0.47}$, we obtain

$$\omega_{\rm rel}/\omega_{\rm rel}(\Delta N = 0) = 1.27^{0.35}_{-0.43} \quad \text{at } 2\sigma\% \text{ l.c.} \tag{6.11}$$

	$\omega_{\rm rel}/\omega_{\rm rel}(\Delta N=0)$	$N_{\rm eff}$
CMB only	$1.50^{+0.90}_{-0.90}$	$0.04\ldots13.37$
CMB + SN-Ia	$1.12^{+0.35}_{-0.42}$	$0.78\ldots6.48$
CMB + PSCz	$1.40^{+0.49}_{-0.56}$	$1.81\ldots9.59$
CMB + σ_8	$1.27^{+0.35}_{-0.43}$	$1.82\ldots7.59$

Table 6.1: Data analysis results: 2σ likelihood intervals on the effective energy density of relativistic particles, $\omega_{\rm rel}/\omega_{\rm rel}(\Delta N=0)$, and on the corresponding effective number of neutrino species, $N_{\rm eff}$, for different data set combinations. Note that the bounds obtained with CMB data only mainly reflect the priors used in the analysis.

Our results are summarized in Table 6.1. Combination of present day CMB data with SN and with LSS data yields a lower bound $N_{\rm eff} > 0.8$ and > 1.8, respectively, with 2σ likelihood content. Our result is in good agreement with the analysis of Hannestad (2001), which considered similar data sets. It is worth emphasizing the fact that $N_{\rm eff} = 0$ is excluded at much more than 2σ: this can be considered as a strong cosmological evidence of the presence of a neutrino background, as predicted by the Standard Model. The upper bounds for the combined sets can be expressed as $N_{\rm eff} < 6.5$ for CMB+SN and $N_{\rm eff} < 9.6$ for CMB+LSS, at 2σ l.c.

6.1.5 Fisher matrix forecast

In this section we perform a Fisher matrix analysis with the technique explained is § 5.2 in order to estimate the precision with which forthcoming satellite experiments will be able to constrain the parameter $z_{\rm eq}$.

Table 6.2 summarizes the experimental parameters for WMAP and Planck employed in the analysis, which considers temperature information only. For both experiment we take a sky coverage $f_{\rm sky} = 0.50$. These values are indicative of the expected performance of the experimental apparatus, but the actual values may be somewhat different, especially for the Planck satellite.

As base parameters for the Fisher matrix analysis, we use the following nine dimensional parameter set:

$$\boldsymbol{\theta} = \left\{\omega_b, \omega_c, \omega_\Lambda, \mathcal{R}^{\rm shift}, z_{\rm eq}, n_s, n_t, r, Q\right\} . \tag{6.12}$$

Here n_s, n_t are the scalar and tensor spectral indices respectively and $r \equiv C_2^T/C_2^S$ is the tensor to scalar ratio at the quadrupole. We adopt a phenomenological normalization parameter, given by

$$Q \equiv \left(\sum_{\ell=2}^{\ell_{\rm max}} \ell(\ell+1)C_\ell\right)^{1/2} , \tag{6.13}$$

so that Q effectively measures the mean power seen by the experiment. The shift parameter $\mathcal{R}^{\rm shift}$, including the radiation content as in Eq. (6.5) takes into account the geometrical degeneracy. Our purely adiabatic reference model has parameters: $\omega_b = 0.0200$ ($\Omega_b = 0.0473$), $\omega_c = 0.1067$ ($\Omega_c = 0.2527$), $\omega_\Lambda = 0.2957$ ($\Omega_\Lambda = 0.7000$), ($h = 0.65$), $\mathcal{R}^{\rm shift} = 0.953$, $z_{\rm eq} = 3045$, $n_s = 1.00$, $n_t = 0.00$, $r = 0.10$, $Q = 1.00$. This is a fiducial, concordance model, which we believe to be in good agreement with most recent determinations of the cosmological

	WMAP			Planck			
ν (GHz)	40	60	90	100	150	220	350
θ_c (degrees)	0.46	0.35	0.21	0.18	0.13	0.09	0.08
$\sigma_c/10^{-6}$	6.6	12.1	25.5	1.7	2.0	4.3	14.4
$w_c^{-1}/10^{-15}$	2.9	5.4	6.8	0.028	0.022	0.047	0.44
ℓ_c	289	385	642	757	1012	1472	1619
$\ell_{\max}$	1500			2000			

Table 6.2: Experimental parameters used in the Fisher matrix analysis for WMAP (first 3 channels) and Planck (last 4 channels).

parameters (flat universe, scale invariant spectral index, BBN compatible baryon content, large cosmological constant). Furthermore, we allow for a modest, 10% tensor contribution at the quadrupole in order to be able to include tensor modes in the Fisher matrix analysis.

We plot the derivatives of C_ℓ with respect to the different parameters in Fig. 6.5. Generally, we note that derivatives with respect to the combination of parameters describing the matter content of the universe (ω_b and ω_c, $\mathcal{R}^{\rm shift}$, $z_{\rm eq}$) are large in the acoustic peaks region, $\ell > 100$, while derivatives with respect to parameters describing the tensor contribution (n_t, r) are important in the large angular scale region. Since measurements in this region are cosmic variance limited, we expect uncertainties in the latter set of parameters to be large regardless of the details of the experiment. The curve for $\partial C_\ell/\partial Q$ is of course identical to the C_ℓ's themselves. The cosmological constant is a notable exception: variation in the value of ω_Λ keeping all other parameters fixed produces a perfect degeneracy in the acoustic peaks region. Therefore we expect the derivative $\partial C_\ell/\partial\omega_\Lambda$ to be zero in this region. Small numerical errors in the computation of the spectra, however, artificially spoil this degeneracy, erroneously leading to smaller predicted uncertainties. In order to suppress this effect, we set $\partial C_\ell/\partial\omega_\Lambda = 0$ for $\ell > 200$. From Eq. (5.76, page 129) we see that a large absolute value of $\partial C_\ell/\partial\theta_i$ leads to a large F_{ii} and therefore to a smaller 1σ error (roughly neglecting non-diagonal contributions). If the derivative along θ_i can be approximated as a linear combination of the others, however, then the corresponding directions in parameter space will be degenerate, and the expected error will be important. This is the case for mild, featureless derivatives such as $\partial C_\ell/\partial r$, while strongly varying derivatives (such as $\partial C_\ell/\partial\mathcal{R}^{\rm shift}$) induce smaller errors in the determination of the corresponding parameter. Therefore the choice of the parameter set is very important in order to correctly predict the standard errors of the experiment.

Error forecast

The quantity $\epsilon_i \equiv 1/\sqrt{\lambda_i}$, where λ_i is the i-th eigenvalue of the Fisher matrix, is sometimes used as a rough indication of the resolving power of an experiment. It expresses the accuracy with which the i-th eigenvector of the Fisher matrix can be determined. The principal components describe to a good approximation which linear combinations of the cosmological parameters can be directly measured with the CMB. In fact, they represent linear approximations to the orthogonal normal parameters introduced in § 4.2. For WMAP (Planck) the

Parameter		WMAP	Planck
Redshift of equality	$\delta z_{\rm eq}/z_{\rm eq}$	0.23	0.02
Relativistic energy	$\delta\omega_{\rm rel}/\omega_{\rm rel}$	0.43	0.03
Effective ν families	$\Delta N_{\rm eff}$	3.17	0.24
Baryons density	$\delta\omega_b/\omega_b$	0.12	< 0.01
CDM density	$\delta\omega_c/\omega_c$	0.50	0.04
Cosmological constant	$\delta\omega_\Lambda/\omega_\Lambda$	3.40	1.71
Shift parameter	$\delta\mathcal{R}^{\rm shift}$	< 0.01	< 0.01
Scalar spectral index	δn_s	0.15	0.01
Tensor spectral index	δn_t	1.96	1.08
Scalar-to-tensor ratio	$\delta r/r$	5.22	2.67
Normalization	δQ	0.01	< 0.01

Table 6.3: Fisher matrix analysis results: expected 1σ errors for the WMAP and Planck satellites. See the text for details and discussion.

number of eigenvectors with $\epsilon_i < 10^{-3}$ is 1 out of 9 (3 out of 9) and with $\epsilon_i < 10^{-2}$ is 3/9 (6/9).

Table 6.3 shows the results of our analysis for the expected 1σ error on the physical parameters. Determination of the redshift of equality can be achieved by WMAP with 23% accuracy, while Planck will pinpoint it down to within 2% relative error. From $\omega_{\rm rel} = (\omega_b+\omega_c)/z_{\rm eq}$ it follows that the energy density of relativistic particles, $\omega_{\rm rel}$, will be determined within 43% by WMAP and 3% by Planck. This translates into an impossibility for WMAP alone of measuring the effective number of relativistic species ($\Delta N_{\rm eff} \approx 3.17$ at 1σ), while Planck will be able to track it down to $\Delta N_{\rm eff} \approx 0.24$. As for the other parameters, while the acoustic peak' positions (through the value of $\mathcal{R}^{\rm shift}$) and the matter content of the universe can be determined by Planck with high accuracy (of the order of or less than one percent), the cosmological constant remains (with CMB data only) almost undetermined, because of the effect of the geometrical degeneracy. One could also see this as a consequence of an inappropriate parameterization of the problem: we should in fact use the parameters which the physics of the CMB measures best, i.e. the principal components. The scalar spectral index n_s and the overall normalization will be well constrained already by WMAP (within 15% and 1%, respectively), while because of the reasons explained above the tensor spectral index n_t and the tensor contribution r will remain largely unconstrained by both experiments. Generally, an improvement of a factor ten is to be expected between WMAP and Planck in the determination of most cosmological parameters.

Our analysis considers temperature information only. Inclusion of polarization measurements would tighten errors, especially for the "primordial" parameters n_s, n_t and r (Zaldarriaga et al., 1997; Bucher et al., 2001). This is especially important for a WMAP-type experiment, since a precise determination of n_s and an higher accuracy in ω_m would greatly improve the precision on $N_{\rm eff}$ which can be obtained with temperature only. By the time Planck will obtain his first results, polarization measurements will hopefully have been performed. Combination of polarization information with the WMAP temperature data would then considerably improve the precision of the extracted parameter values.

A Fisher matrix analysis for $\Delta N_{\rm eff}$ was previously performed by Lopez et al. (1999)

and repeated by Kinney & Riotto (1999) with the equivalent chemical potential ξ, $\Delta N = 15/7(2(\xi/\pi)^2 + (\xi/\pi)^4)$, and a strong degeneracy was found between $N_{\rm eff}, h$ and Ω_Λ, and to lesser extent with Ω_b. We have seen here that the degeneracy really is between $\omega_{\rm rel}, \omega_m$ and n_s, and the degeneracy previously observed is thus explained because they considered flat models, where a change in Ω_Λ is equivalent to a change in ω_m, $\omega_m = (1 - \Omega_\Lambda - \Omega_b)h^2$. The results regarding how precisely the future satellite missions can extract the relativistic energy density, can be translated into approximately $\Delta N_{\rm eff} = 3.17$ ($\xi = 2.4$) and $\Delta N_{\rm eff} = 0.24$ ($\xi = 0.73$) for WMAP and Planck respectively. However, including neutrino oscillation leads to equilibration of the different chemical potentials, and hence BBN leads to the stronger bound $|\xi| < 0.07$ for all neutrino species (Dolgov et al., 2002).

Comparison with WMAP data analysis

After the release of the WMAP first year observations, several groups have independently carried out an analysis similar to the one presented above (Crotty et al., 2003b; Hannestad, 2003; Pierpaoli, 2003). Unfortunately, none of these works includes tensor modes as in our forecasts, and one has to keep in mind that the FMA assumed temperature information only and experimental parameters as appropriate for the original mission specifications, which may be slightly different from the effective parameters for the first year only. Despite the fact that the details of the data included and the prior assumptions vary for each work, the overall agreement of their findings with our forecasts for WMAP is nonetheless very satisfactory. We briefly review their conclusions and compare them with the above predictions.

In Crotty et al. (2003b) the 1σ error on $N_{\rm eff}$ is found to be $\Delta N_{\rm eff} = 3.4$ using WMAP data only (but including the TE-spectrum) and a weak top-hat prior on the Hubble parameter, $0.5 < h < 0.9$, with the analysis limited to flat models only. This result has to be contrasted with the prediction above, which for the full WMAP data gives (at 1σ) $\Delta N_{\rm eff} = 3.17$. As predicted, the WMAP observations improve dramatically on the bounds for $N_{\rm eff}$ from CMB only, which become with the above assumptions $-2.1 < \Delta N_{\rm eff} < 6.9$ (at 2σ likelihood content).

These findings are in good concordance with the more general set-up of Pierpaoli (2003), where curved models are considered as well, the CBI data are used together with the WMAP observations and constraints from the 2dF matter power spectrum are also included. In this case the results do not compare directly with our predictions because of the inclusion of external constraints in the form of the matter power spectrum. The 95% likelihood interval is then tighter because of the more powerful observational data used, giving (without Hubble prior) $\Delta N_{\rm eff} = 5.5$.

The quite complete investigation of Hannestad (2003) also derives constraints on the neutrino masses, and considers the effects of the inclusion of further observational constraints, such as a prior on the Hubble parameter, a prior on Ω_m from supernovæ data, a BBN prior on ω_b and the 2dF matter power spectrum. Where comparable, the findings are entirely compatible with the other two works; in particular, for the case of massless neutrinos and WMAP data only, the 95% likelihood interval for flat models only and a weak top-hat prior $0.5 \leq h \leq 0.85$ is $\Delta N_{\rm eff} = 8.9$.

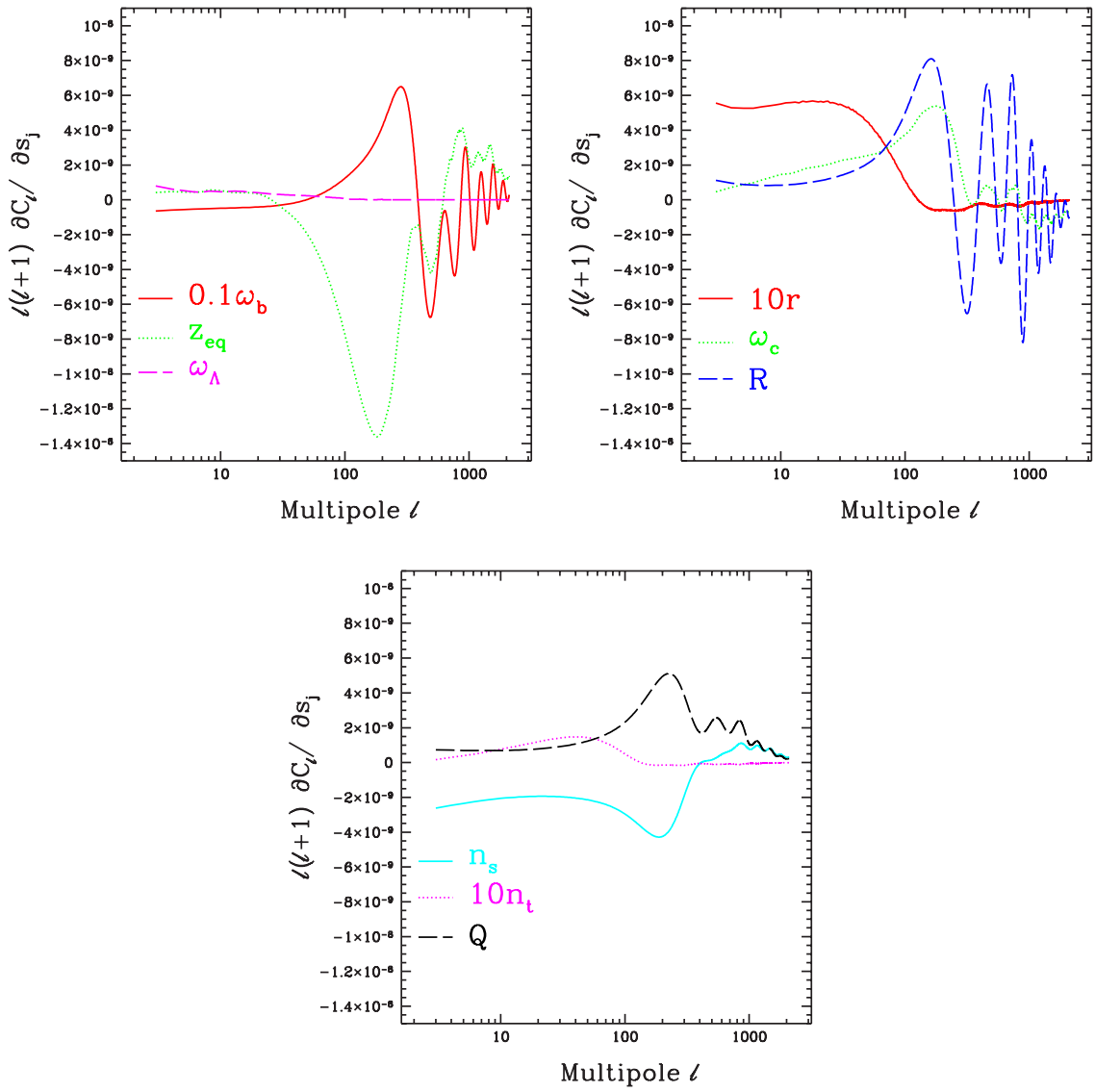

Figure 6.5: Derivatives of C_ℓ with respect to the 9 parameters evaluated at the reference model described in the text. The numerical prefactor indicates that the corresponding curve has been rescaled: thus $0.1\omega_b$ means that the displayed curve is $0.1 \cdot \partial C_\ell / \partial \omega_b$. The derivative $\partial C_\ell / \partial \omega_\Lambda$ has been set to 0 for $\ell > 200$ in order to suppress the effect of numerical errors, thus taking into account the geometrical degeneracy.

6.2 The primordial helium fraction

This section is based on the work Trotta & Hansen (2004), where the first determination of the helium abundance from CMB data alone was presented. After giving the motivation underlying this investigation in § 6.2.1, we discuss in § 6.2.2 the role of the helium mass fraction for CMB anisotropies, and in particular the details of the ionization history of the Universe which are relevant for constraining the helium abundance with the CMB. We then review the standard Big-Bang Nucleosynthesis scenario for the abundance of light elements and compare its predictions with current astrophysical measurements in § 6.2.3; the present constraints from CMB data are presented in § 6.2.4, while the future potential of using the CMB as an independent way of determining the helium abundance is elucidated in § 6.2.5. There we also explore the impact of helium for future accurate determination of the baryon abundance.

6.2.1 Motivation

Our understanding of the baryon abundance has increased dramatically over the last few years, coming from two independent paths, namely BBN and CMB. Absorption features from high-redshift quasars allow us to measure precisely the deuterium abundance, D/H, which combined with BBN calculations provides a reliable estimate of the baryon to photon ratio,

$$\eta_{10} \equiv \frac{n_b}{n_\gamma} 10^{10} \,. \tag{6.14}$$

An independent determination of the baryon content of the universe from CMB anisotropies comes from the increasingly precise measurements of the acoustic peaks, via the characteristic signature of the photon-baryon fluid oscillations discussed in § 4.1.2.2. The agreement between these two completely different approaches is both remarkable and impressive (see details below). The time is therefore ripe to proceed and test the agreement between other light elements which are also probed both by BBN and CMB.

Helium being the most abundant of the light elements, it is natural to focus on this element by exploring the dependency of CMB anisotropies on the value of the primordial helium mass fraction Y_p, defined as

$$Y_p \equiv 4 \frac{n_{\rm He}}{n_b} \,, \tag{6.15}$$

where $n_{\rm He}$ and n_b denote the number densities of ^{4}He atoms and baryons, respectively. If we denote by n_N and n_P the number densities of neutrons and protons, respectively, and assume that all neutrons are in He nuclei, then a simple counting argument gives the estimate

$$Y_p = \frac{2 n_N/n_P}{1 + n_N/n_P} \approx 0.25 \,, \tag{6.16}$$

where the numerical value comes from a rough approximation to the freeze-out value of the neutron to proton ratio $n_N/n_P \approx 1/7$, see e.g. Kolb & Turner (1990). The detailed value of Y_p is predicted by BBN as a function of two parameters only, the baryon abundance and the number of relativistic degrees of freedom at BBN (Fields & Sarkar, 2004).

The hope is that the CMB observations might provide an independent measurement of Y_p, accurate enough to help clarify the present-day discrepancies between direct observations

of the helium fraction derived from astrophysical systems, whose errors are seemingly dominated by systematics which are hard to assess. The latest CMB data are precise enough to allow taking this further step, and in view of the emerging "baryon tension" between BBN predictions from observations of different light elements (Cyburt et al., 2003) possibly requires taking such a step. The advantage of using CMB anisotropies rather than the traditional astrophysical measurements, is that the CMB provides a clear measurement of the primordial helium fraction before it could be changed by any astrophysical process. On the other hand the dependence of the CMB power spectrum on the primordial helium fraction is rather mild, a fact which makes it presently safe to set the helium mass fraction to a constant for the purpose of CMB data analysis of other cosmological parameters, but will have an impact on the baryon abundance determination from Planck quality data, as we show in § 6.2.5.

6.2.2 The impact of helium on the CMB: ionization history revisited

We now resume our discussion of the recombination epoch and reionization history of the Universe sketched in § 4.1.3, and focus on the role of the helium mass fraction, considered here as a free parameter. In a second step, the aim will be to combine the CMB results with the BBN predictions and compare the result with the independent astrophysical determinations of the light elements abundance. We thus have at our disposal three different tools, each of which probes the same quantities at three vastly different epochs of the cosmic history. It is important to stress that a good agreement among the three is by no means trivial, and that testing their concordance is a powerful way to check the consistency of the standard cosmological scenario. On the other hand, significant discrepancies would necessary imply the need for new physics.

The recent WMAP data allow us to determine with very high precision the epoch of photon decoupling, $z_{\rm dec}$, i.e. the epoch at which the ionized electron fraction, $x_e(z) = n_e/n_H$, has dropped from 1 to its residual value of order 10^{-4}. Here n_e denotes the number density of free electrons, while n_H is the total number density of H atoms (both ionized and recombined). The redshift of decoupling has been determined to be $z_{\rm dec} = 1088^{+1}_{-2}$ (Spergel et al., 2003), which corresponds to a temperature of about 0.25 eV. Helium recombines earlier than hydrogen, roughly in two steps: around redshift $z = 6000$ HeIII recombines to HeII, while HeII to HeI recombination begins around $z < 2500$ and finishes just after the start of H recombination (Libarskii & Sunyaev, 1983; Hu et al., 1995; Seager et al., 1999, 2000).

The baryon number density per m^3 $n_b(z)$ is related to the baryon energy density today, ω_b, by

$$n_b = 11.3(1+z)^3\omega_b \tag{6.17}$$

and we have $n_H = n_b(1 - Y_p)$. Usually, the ionization history is described in terms of $x_e(z) = n_e/(n_b(1 - Y_p))$. However, for the purpose of discussing the role of Y_p, it is more convenient to consider the quantity

$$f_e(z) \equiv n_e/n_b \tag{6.18}$$

instead, the ratio of free electrons to the total number of baryons. For brevity, we will call f_e the free electron fraction. Once the baryon number density has been set by fixing ω_b,

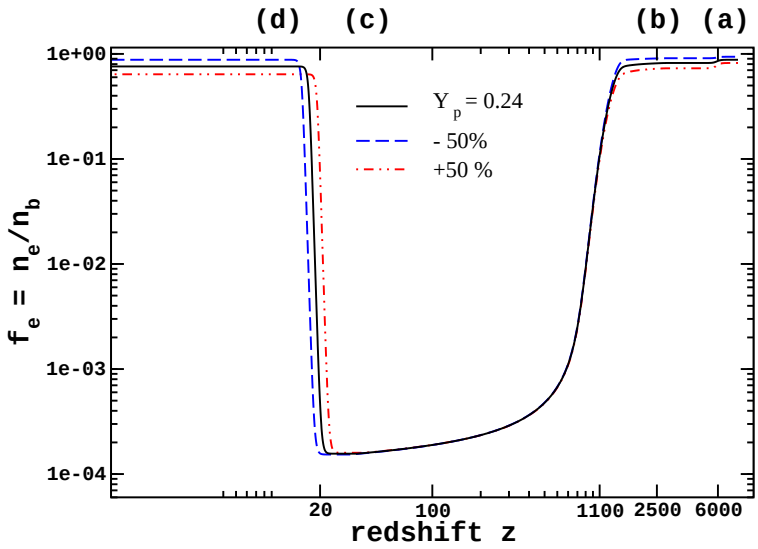

Figure 6.6: Evolution of the number density of electrons normalized to the number density of baryons, $f_e = n_e/n_b$, as a function of redshift for different values of the helium fraction Y_p. The black-solid curve corresponds to the standard value $Y_p = 0.24$. The labels (a) to (d) indicate the four different phases discussed in the text.

one can think of Y_p as an additional parameter which controls the number of free electrons available in the tight coupling regime. The CMB power spectrum depends on the full detailed evolution of the free electron fraction, but we can qualitatively describe the role of helium in four different phases of the ionization/recombination history, displayed in Fig. 6.6.

(a) Before HeIII recombination, all electrons are free, therefore $f_e(z > 6000) = 1 - Y_p/2$.

(b) HeII progressively recombines and just before H recombination begins, f_e has dropped to the value $f_e(z \approx 1100) = 1 - Y_p$.

(c) After decoupling, a residual fraction of free electrons freezes out, giving $f_e(30 \lesssim z \lesssim 800) = f_e^{\rm res} \approx 2.7 \cdot 10^{-5}\sqrt{\omega_m}/\omega_b$.

(d) Reionization of all the H atoms gives $f_e(z \lesssim 20) = 1 - Y_p$.

During phase (a), the photon-baryons fluid is in the tight coupling regime. However the presence of ionized He increases diffusion damping, therefore having an impact on the damping scale in the acoustic peaks region: the diffusion damping length (3.120, page 78) including helium can be approximated as (Hu & Sugiyama, 1995a)

$$\lambda_{\rm D}^2 \approx 1.7 \times 10^7 \left(1 - \frac{Y_p}{2}\right)^{-1} \omega_b^{-1} \omega_m^{-1/2} a^{5/2} \frac{1}{3\sqrt{a_{\rm eq}/a} + 2} \quad {\rm Mpc}^2 \,. \tag{6.19}$$

As expected, a larger helium fraction implies an increased damping length, and thus an extra power suppression on small scales.

When the detailed energy level structure of HeII is taken into account (Seager et al., 2000), the transition to phase (b) is smoother than in the Saha equation approximation. Therefore the plateau with $f_e = 1 - Y_p$ is not visible in Fig. 6.6. Before H recombination, He atoms remain tightly coupled to H atoms through collisions, with the same dynamical behavior. In particular, it is the total ω_b which determines the amount of gravitational pressure on the photon-baryons fluid, and which sets the acoustic peak enhancement/suppression, see § 4.1.2. Hence we do not expect the value of Y_p to have any influence on the boosting (suppression) of odd (even) peaks. The redshift of decoupling (transition between (b) and (c)) depends mildly on Y_p in a correlated way with ω_b, since the number density of free electrons in the tight coupling regime (just before H recombination) scales as $n_e = f_e n_b = n_b(1 - Y_p)$. Hence an increase in ω_b can be compensated by allowing for a larger helium fraction. An analytical estimate along the same lines as in e.g. Kolb & Turner (1990) indicates that a 10% change in Y_p affects $z_{\rm dec}$ by roughly 0.1%, which corresponds to $\Delta z_{\rm dec} \approx 1$. This is of the same order as the current 1σ errors on $z_{\rm dec}$, obtained by fixing $Y_p = 0.24$.

After H recombination, the residual ionized electron fraction $f_e^{\rm res}$ does not depend on Y_p, but is inversely proportional to the total baryon density (phase (c)). As the CMB photons propagate, they are occasionally rescattered by the residual free electrons. The corresponding optical depth, $\tau^{\rm res}$ is given by

$$\begin{aligned}\tau^{\rm res} &= \int_{t_0}^{t_{\rm dec}} n_e^{\rm res} c\sigma_T {\rm d}t \\ &\approx 1.86 \cdot 10^{-6} \sqrt{\Omega_m} \int_0^{z_{\rm dec}} \frac{(1+z)^2}{[(1+z)^3 + \Omega_\Lambda/\Omega_m]^{1/2}} dz\,.\end{aligned} \tag{6.20}$$

Performing the integral we can safely neglect the contribution of the cosmological constant at small redshift, since $z_{\rm dec} \gg \Omega_\Lambda/\Omega_m$. Retaining only the leading term, the approximated optical depth from the residual ionization fraction is estimated to be

$$\tau^{\rm res} \approx 1.24 \cdot 10^{-6} (1 + z_{\rm dec})^{3/2} \approx 0.045, \tag{6.21}$$

independent of the cosmological parameters and of the helium fraction. Therefore after last scattering we do not expect any significant effect on CMB anisotropies coming from the primordial helium fraction, until the reionization epoch, phase (d).

As pointed out in § 4.1.3.2, CMB anisotropies are sensitive only to the integrated reionized fraction if temperature information only is available, while specific signatures are imprinted on the E-polarization and ET-cross correlation power spectra by the detailed shape of the reionization history. There are several physically motivated reionization scenarios, which however cannot be clearly distinguished at present (Haiman & Holder, 2003; Hansen & Haiman, 2004). Therefore at the present level of accuracy it is safe for our purpose to assume an abrupt reionization, i.e. that at the reionization redshift $z_{\rm re}$ all the hydrogen was quickly reionized, thus producing a sharp rise of n_e from its residual value to n_H. More precisely, $z_{\rm re}$ is the redshift at which $x_e(z_{\rm re}) = 0.5$. In our treatment we neglect HeII reionization, for which there is evidence at a redshift $z \approx 3$ (see Theuns et al., 2002 and references therein). This effect is small, since one extra electron released at $z \approx 3$ would change the reionization optical depth only by about 1%. The effect of HeIII reionization, which happens still later, is even smaller. We also neglect the increase of the helium fraction due to non-primordial

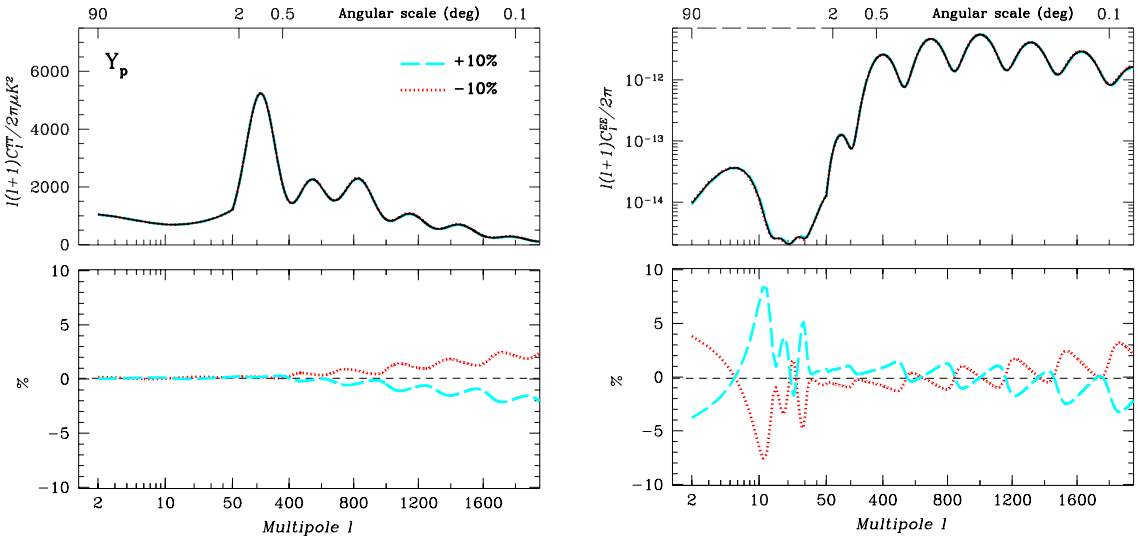

Figure 6.7: CMB temperature (left panel) and polarization (right panel) power spectra and percentage change (bottom panels) for a 10% larger (smaller) value of the helium mass fraction, Y_p. The solid-black line in the top panels corresponds to a standard ΛCDM model, with $Y_p = 0.24$. The impact is at the percent level, and is almost indistinguishable in the top panels. All other parameters are fixed to the value of our fiducial model (Table 6.4), in particular, we have $\tau_{\rm re} = 0.166$.

helium production, which has a negligible effect on CMB anisotropies. Those approximations do not affect the results at today's level of sensitivity of CMB data: for WMAP noise levels, even inclusion of the polarization spectra is not enough to distinguish between a sudden reionization scenario and a more complex reionization history. At the level of Planck a more refined modelling of the reionization mechanism will be necessary (Holder et al., 2003; Doroshkevich et al., 2003).

In the sudden reionization scenario adopted here, the relation between reionization redshift and reionization optical depth, $\tau_{\rm re}$, is given by Eq. (4.38, page 99). Once again, since the number density of reionized electrons scales as $\omega_b(1 - Y_p)$, the redshift of reionization is positively correlated with Y_p (for fixed optical depth and baryon density).

As a result of the physical mechanism described above, a 10% change in Y_p has a net impact on the CMB power spectrum at the percent level. The impact on the CMB temperature and polarization power spectra is highlighted in Fig. 6.7. In the temperature panel, we notice that a larger helium fraction slightly suppresses the peaks because of diffusion damping, while it has no impact on large scales. Polarization is induced by the temperature quadrupole component at last scattering and the reionization bump induced in the polarization spectrum (see § 4.1.3.2) is clearly visible in the polarization panel of Fig. 6.7 in the $\ell \approx 15$ region. A change in the helium fraction implies a shift of the redshift of reionization for a given (fixed) optical depth, and a consequent shift of the position of the reionization bump via Eq. (4.41, page 99). The value of Y_p does not affect the height of the bump, which is controlled by the optical depth and is proportional to τ^2. This effect is highlighted in the

polarization panel of Fig. 6.7: a 10% change in Y_p induces roughly a 10% change in the position of the bump. The subsequent two oscillatory features for $\ell \lesssim 50$ reflect the displacement of further secondary reionization induced polarization oscillations. However, since the value of polarized power is very low in that region, such secondary oscillations are very hard to detect precisely.

In principle, given an accurate knowledge of the reionization history, the effect of Y_p on the polarization bump would assist in the determination of the helium abundance. However, our ignorance of the reionization history prevents us from recovering useful information out of the measured reionization bump. The displacement induced by Y_p is in fact degenerate with a partial reionization, or with other, more complex reionization mechanisms. Hence constraints on Y_p come effectively from the damping tail in the $\ell \gtrsim 400$ region of the temperature spectrum, which needs to be measured with very high accuracy. Other light elements like deuterium and helium-3 are much less abundant, and will therefore have even smaller effect on the CMB power spectrum, at the order of 10^{-5}.

6.2.3 Astrophysical measurements and BBN predictions

Once we fix the number of relativistic degrees of freedom by specifying the number of massless neutrino families, the standard model of Big-Bang Nucleosynthesis (BBN) has only one free parameter, namely the baryon to photon ratio η_{10} defined in (1.47, page 31), which for long has been known to be in the range $1 - 10$ (Kolb & Turner, 1990). Thus by observing just one primordial light element one can predict the abundances of all the other light elements.

Astrophysical measurements

The deuterium to hydrogen abundance, D/H, is observed by Ly-α features in several quasar absorption systems at high redshift, $D/H = 2.78^{+0.44}_{-0.38} \times 10^{-5}$ (Kirkman et al., 2003), which in BBN translates into the baryon abundance, $\eta_{10} = 5.9 \pm 0.5$. Using BBN one thus predicts the helium mass fraction to be in the range $0.2470 < Y_p < 0.2487$. The dispersion in various deuterium observations is, however, still rather large, ranging from $D/H = 1.65 \pm 0.35 \times 10^{-5}$ (Pettini & Bowen, 2001) to $D/H = 3.98^{+0.59}_{-0.67} \times 10^{-5}$ (Kirkman et al., 2003), which most probably indicates underestimated systematic errors.

The observed helium mass fraction comes from the study of extragalactic HII regions in blue compact galaxies. A careful study by Izotov & Thuan (1998) gives the value $Y_P = 0.244 \pm 0.002$; however, also here there is a large scatter in the various observed values, ranging from $Y_p = 0.230 \pm 0.003$ (Olive et al., 1997) over $Y_p = 0.2384 \pm 0.0025$ (Peimbert et al., 2002) and $Y_p = 0.2391 \pm 0.0020$ (Luridiana et al., 2003) to $Y_p = 0.2452 \pm 0.0015$ (Izotov et al., 1999). Besides the large scatter there is also the problem that the helium mass fraction predicted from observations of deuterium combined with BBN, $0.2470 < Y_p < 0.2487$, is larger than (and seems almost in disagreement with) most of the observed helium abundances, which probably points towards underestimated systematic errors, rather than the need for new physics (Cyburt et al., 2003; Barger et al., 2003b). Figure 6.8 is a compilation of the above measurements, and offers a direct comparison with the current (large) errors from CMB observations, presented in § 6.2.4, and with the potential of future CMB measurements, discussed in § 6.2.5.

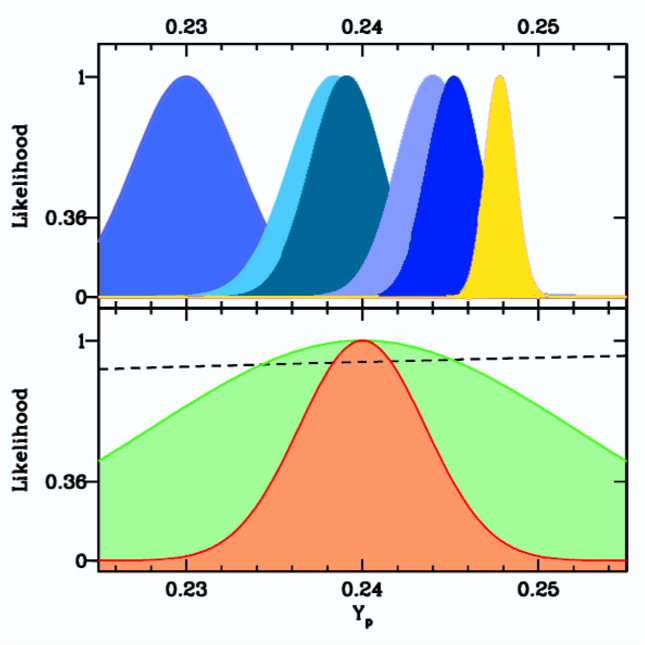

Figure 6.8: In the top panel we plot a few current direct astrophysical measurements of the helium mass fraction Y_p as Gaussian likelihood curves with standard deviation corresponding to the given 1σ (statistical) error (blue/dark gray curves, on the left of the diagram), and the value inferred from deuterium measurements combined with BBN (yellow/light gray curve, on the far right), see the text for references. In the bottom panel, a direct comparison with CMB present-day accuracy (actual CMB data, black dashed line, this work; the 1σ likelihood interval is $0.16 < Y_p < 0.50$) and with its future potential (Fisher matrix forecast for Planck – green/light gray curve – and a Cosmic Variance Limited experiment – orange/dark gray curve).

The observed abundance of primordial ^{7}Li using the Spite plateau is possibly spoiled by various systematic effects (Ryan et al., 2000; Salaris & Weiss, 2001). Therefore it is more appropriate to use the BBN predictions together with observations to estimate the depletion factor $f_7 = {}^7\mathrm{Li}_{\mathrm{obs}}/{}^7\mathrm{Li}_{\mathrm{prim}}$ instead of using $^7\mathrm{Li}_{\mathrm{obs}}$ to infer the value of η_{10} (Burles et al., 2001; Hansen et al., 2002).

The numerical predictions of standard BBN (as well as various non-standard scenarios) have reached a high level of accuracy (Lopez & Turner, 1999; Esposito et al., 2000a,b; Burles et al., 2001), and the precision of these codes is well beyond the systematic errors discussed above.

BBN and the need for new physics

If the CMB-determined helium mass fraction turns out to be as high as suggested by BBN calculations together with the CMB observation of $\Omega_b h^2$ (as discussed above), this could indicate a systematic error in the present direct astrophysical helium observations.

Alternatively, if the CMB could independently determine the helium value with sufficient precision to confirm the present helium observations, then this would be a smoking gun for

new physics. In fact, one could easily imagine non-standard BBN scenarios which would agree with present observations of η_{10}, while having a low helium mass fraction. All what is needed is additional non-equilibrium electron neutrinos produced at the time of neutrino decoupling which would alter the $n-p$ reaction. This could alter the resulting helium mass fraction while leaving the deuterium abundance unchanged. One such possibility would be a heavy sterile neutrino whose decay products include ν_e. A sterile neutrino with life-time of $1-5$ sec and with decay channel $\nu_s \to \nu_e + \phi$ with ϕ a light scalar (like a majoron), would leave the deuterium abundance roughly untouched, but can change the helium mass fraction between $\Delta Y_p = -0.025$ and $\Delta Y_p = 0.015$ if the sterile neutrino mass is in the range $1-20$ MeV (Dolgov et al., 1999). A simpler model would be standard neutrino oscillation between a sterile neutrino and the electron neutrino. The lifetime is about 1 sec when the sterile state has mass about 10 MeV, and the decay channel is $\nu_s \to \nu_e + l + \bar{l}$ (with l any light lepton), and such masses and life-times are still unconstrained for large mixing angle (Dolgov et al., 2000). Related BBN issues are discussed by Shi et al. (1999); Di Bari & Foot (2001); Kirilova (2003). Such possibilities are hard to constrain without an independent measurement of the helium mass fraction.

Another much studied effect of neutrinos is the increased expansion rate of the universe if additional degrees of freedom are present (for BBN), and the degeneracy between the total density in matter and relativistic particles (for CMB), which is presented in detail in § 6.1. The more general set-up would then be to allow $N_{\rm eff}$ as a further free parameter both in the CMB and BBN analysis, but because of the very weak dependence of the CMB on Y_p this would spoil any hope of being able to constrain the helium fraction with the CMB; therefore we choose to fix $N_{\rm eff} = 3.04$.

Also, an electron neutrino chemical potential could potentially alter the BBN predictions (Kang & Steigman, 1992; Lesgourgues & Pastor, 1999), however, with the observed neutrino oscillation parameters the different neutrino chemical potentials would equilibrate before the onset of BBN (Dolgov et al., 2002; Wong, 2002; Abazajian et al., 2002), hence virtually excluding this possibility (see however Barger et al., 2003a).

6.2.4 WMAP Monte Carlo analysis

We use a modified version of the publicly available Markov Chain Monte Carlo package COSMOMC as described in Lewis & Bridle (2002) in order to construct Markov chains (see § 5.1.7) in our seven dimensional parameter space. We sample over the following set of cosmological parameters: the physical baryon and CDM densities, $\omega_b \equiv \Omega_b h^2$ and $\omega_c \equiv \Omega_c h^2$, the cosmological constant in units of the critical density, Ω_Λ, the scalar spectral index and the overall normalization of the adiabatic power spectrum, n_s and $A_s \equiv \zeta_0^2$, cf. Eq. (4.6, page 91), the redshift at which the reionization fraction is a half, $z_{\rm re}$, and the primordial helium mass fraction, Y_p. We restrict our analysis to flat models, therefore the Hubble parameter is a derived parameter,

$$h = [(\omega_c + \omega_b)/(1-\Omega_\Lambda)]^{1/2}\,. \tag{6.22}$$

We consider purely adiabatic initial conditions and three massless neutrino families for the reason given above. We do not consider either gravitational waves or massive neutrinos. We include the WMAP data from Kogut et al. (2003); Hinshaw et al. (2003b) (temperature and polarization) with the routine for computing the likelihood supplied by the WMAP team

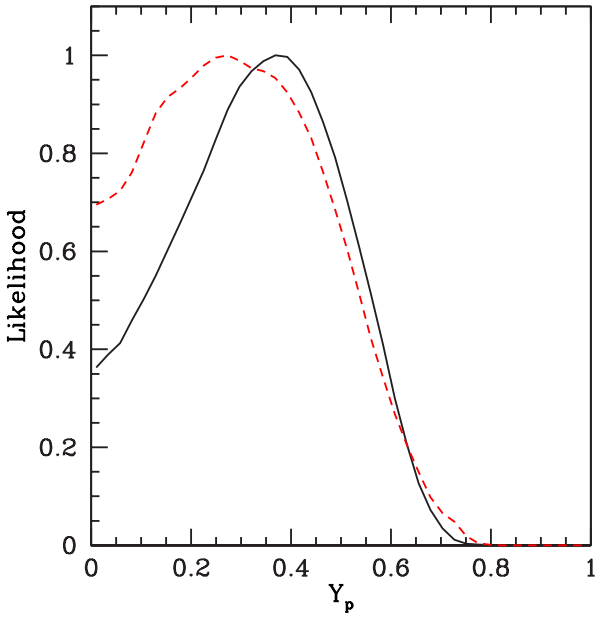

Figure 6.9: One-dimensional posterior likelihood distribution for the helium mass fraction, Y_p, using CMB data only. The solid-black line is for all other parameters marginalized, the dashed-red line gives the mean likelihood.

(Verde et al., 2003). We make use of the CBI (Pearson et al., 2003) and of the decorrelated ACBAR (Kuo et al., 2004) band powers above $\ell = 800$ to cover the small angular scale region of the power spectrum.

Since Y_p is a rather flat direction in parameter space with present-day data, we find that a much larger number of samples is needed in order to achieve good mixing and convergence of the chains in the full 7D space. We use $M = 4$ chains, each containing approximately $N = 3 \cdot 10^5$ samples. The mixing diagnostic is carried out along the same lines as in Verde et al. (2003), by means of the Gelman and Rubin criterion (Gelman & Rubin, 1992b). The burn-in of the chains also takes longer than in the case where Y_p is held fixed, and we discard 6000 samples per chain.

Results

Marginalizing over all other parameters, we find that the helium mass fraction from CMB alone is constrained to be

$$Y_p < 0.647 \quad \text{at } 99\% \text{ l.c. (1 tail limit)} \tag{6.23}$$

$$\text{and} \quad 0.160 < Y_p < 0.501 \quad \text{at } 68\% \text{ l.c. (2 tails).} \tag{6.24}$$

Thus, for the first time the primordial helium mass fraction has been observed using the cosmic microwave background. However, present-day CMB data do not have by far sufficient resolution to discriminate between the astrophysical helium measurements, $Y_p \sim 0.244$, and the deuterium guided BBN predictions, $Y_p \sim 0.248$, which would require percent precision.

In Fig. 6.9 we plot the marginalized and the mean likelihood of the Monte Carlo samples as a function of Y_p. If the likelihood distribution is Gaussian, then the 2 curves should be indistinguishable. The difference between marginalized and mean likelihood for Y_p indicates that the marginalized parameters are skewing the distribution, and therefore that correlations play an important role. Although the mean of the 1D marginalized likelihood is rather high, $\langle L(Y_p)\rangle = 0.33$, the mean likelihood peaks in the region indicated by astrophysical measurements, $Y_p \sim 0.25$. In view of this difference, it is important to understand the role of correlations with other parameters, and we will turn to this issue now.

In Fig. 6.10 we plot joint 68% and 99% confidence contours in the (ω_b, Y_p)-space. From the Monte Carlo samples we obtain a small and negative correlation coefficient between the two parameters, $\mathrm{corr}(Y_p, \omega_b) = -0.14$. Baryons and helium appear to be anticorrelated simply because present-day WMAP data do not map the peaks structure to sufficiently high ℓ. Precise measurements in the small angular scale region should reveal the expected positive correlation between the baryon and helium abundances, which is potentially important in order to correctly combine BBN predictions and CMB measurements of the baryon abundance. We turn to this question in more detail in the next section. In BBN the baryon fraction and helium fraction are correlated along a different direction, cf. Fig. 6.10. However, this correlation is very weak, and the BBN relation gives practically a flat line. Since the two parameters are not independent from the CMB point of view, it is in fact not completely accurate to perform the CMB analysis with fixed helium mass fraction of $Y_p = 0.24$ to get the error-bars on the baryon fraction, and then re-input this baryon fraction (and error-bars) to predict the helium mass fraction from BBN. The most accurate procedure is to analyse the CMB data leaving Y_p as a free parameter, thereby obtaining the correct (potentially larger) error-bars on ω_b upon marginalization over Y_p.

In view of the emerging baryon tension between CMB and BBN, it is important to check whether allowing helium as a free parameter can significantly change the CMB determination of the baryon density or its error. In order to evaluate in detail the impact of Y_p on the error-bars for ω_b, we consider the following three cases.

(a) The usual case, when the helium fraction for the CMB analysis is assumed to be known *a priori* and is fixed to the canonical value $Y_p = 0.24$.

(b) A case with a weak astrophysical Gaussian prior on the helium fraction, which we take to be $Y_p = 0.24 \pm 0.01$. As discussed above, the error-bars of the astrophysical measurements are typically a factor 5 tighter than this, but our prior is chosen to encompass the systematic spread between the different observations.

(c) The case in which we assume a uniform prior for Y_p in the range $0 \leq Y_p \leq 1$, i.e. Y_p is considered as a totally free parameter.

We do not find any significant change in the error-bars for ω_b in the three different cases. The confidence intervals on ω_b alone are determined to be (case (c)) $0.0221 < \omega_b < 0.0245$ at 68% l.c. ($0.0204 < \omega_b < 0.0276$ at 99 % l.c.). The standard deviation of ω_b as estimated from the Monte Carlo samples is found to be $\hat{\sigma}_b = 1.3 \cdot 10^{-3}$. This is in complete agreement with the error-bars on ω_b obtained by the WMAP team for the standard ΛCDM case (Spergel et al., 2003). We conclude that at the level of precision of present-day CMB data, it is still

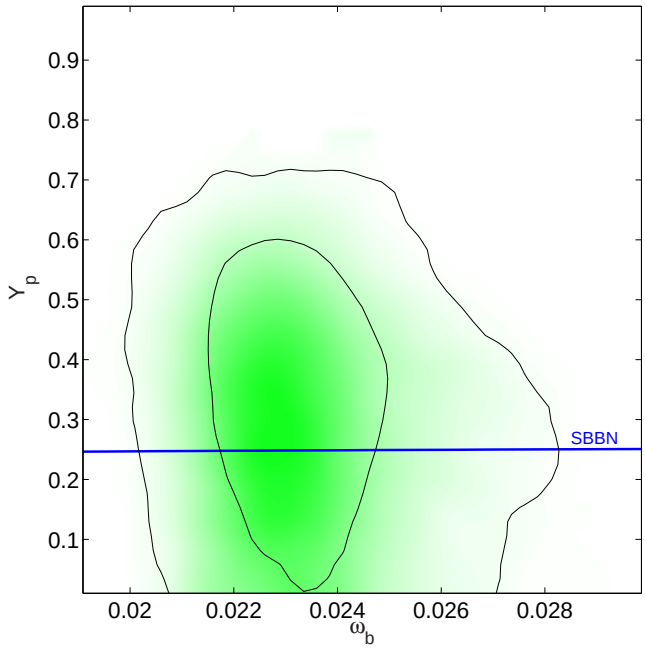

Figure 6.10: Joint 68% and 99% likelihood contours in the (ω_b, Y_p)-plane from CMB data alone. The solid-blue line gives the BBN prediction (Burles et al., 1999), which on this figure almost looks like a straight line.

safe to treat the baryon abundance and the helium mass fraction as independent parameters. This result is non-trivial, since the fact that the damping tail is not yet precisely measured above the second peak would a priori suggest that degeneracies between Y_p, ω_b and n_s could potentially play a role once the assumption of zero uncertainty on Y_p is relaxed. The impact of Y_p is small enough, and the error-bars on ω_b large enough that a uniform prior on Y_p can still be accommodated within the uncertainty in the baryon abundance obtained for case (a). However, the $Y_p - \omega_b$ correlation will have to be taken into account to correctly analyze future CMB data, with a quality such as Planck. We discuss this potential in the next section.

We observe the expected correlation between the redshift of reionization and the helium fraction (Fig. 6.11), which is discussed above. The correlation coefficient between the two parameters is found to be rather large and positive, $\mathrm{corr}(Y_p, z_{\rm re}) = 0.40$. This correlation produces a noticeable change in the marginalized 1D-likelihood distribution for $z_{\rm re}$ as we go from case (a) to case (c). Marginalization over the additional degree of freedom given by Y_p broadens considerably the error-bars on $z_{\rm re}$. In fact, the 68% confidence interval for $z_{\rm re}$ increases by roughly 20% (and shifts to somewhat higher values), from $10.2 - 20.9$ (case (a)) to $10.6 - 23.3$ (case (c)). Case (b) exhibits similar error-bars as case (a). On the other hand, the determination of the reionization optical depth is not affected by the inclusion of helium as a free parameter, giving in all cases $0.08 < \tau_{\rm re} < 0.23$. Correspondingly, the correlation is less significant, $\mathrm{corr}(Y_p, \tau_{\rm re}) = -0.11$. We therefore conclude that the differences in the determination of $z_{\rm re}$ are due only to the variation of the amount of electrons available for reionization as Y_p is changed.

Leaving Y_p as a free parameter also has an impact on the relation between ω_b and the

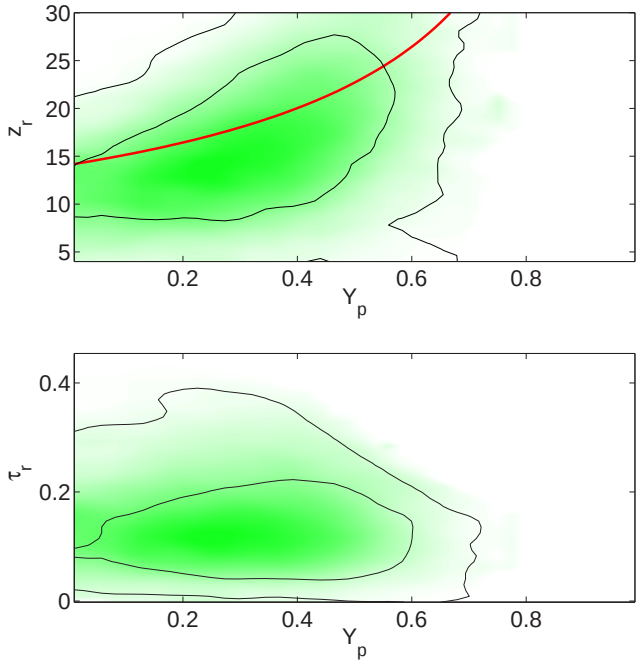

Figure 6.11: Joint 68% and 99% likelihood contours in the $(Y_p, z_{\rm re})$-plane (upper panel) and in the corresponding $(Y_p, \tau_{\rm re})$-plane (bottom panel) from CMB data alone. In the upper panel, the solid-red line is the relation $z_{\rm re}(Y_p)$ from Eq. (4.38, page 99), obtained by fixing the reionization optical depth to the value $\tau_{\rm re} = 0.166$, while the other parameters are those of our fiducial ΛCDM model. Although clearly the exact shape of $z_{\rm re}(Y_p)$ depends on the particular choice of cosmology, it is apparent that the $Y_p - z_{\rm re}$ degeneracy is along this direction. The correlation between $Y_p - \tau_{\rm re}$ is almost negligible with present-day data (bottom panel).

scalar spectral index, n_s. The extra power suppression on small scales which is produced by a larger Y_p can be compensated by a blue spectral index, cf. Fig. 6.12.

6.2.5 Potential of future CMB observations

In order to estimate the precision with which future satellite CMB measurements will be able to constrain the helium mass fraction we perform a Fisher matrix analysis along the lines presented in § 5.2. As already emphasized, in order to obtain a reliable prediction, it is extremely important to choose a parameter set whose effect on the CMB power spectrum is as linear and uncorrelated as possible. Here we improve upon the choice made in § 6.1.5 by adopting the full set of normal parameters introduced in § 4.2. Our nine dimensional basis parameter set is then

$$\boldsymbol{\theta} = \{\mathcal{A}, \mathcal{B}, \mathcal{V}, \mathcal{R}, \mathcal{M}, \mathcal{T}, A_s, n_s, Y_p\} \ , \tag{6.25}$$

where the scalar power spectrum normalization constant is $A_s = \zeta_0^2$, see (4.6, page 91). The quantities $\mathcal{A}, \mathcal{B}, \mathcal{V}, \mathcal{R}, \mathcal{M}, \mathcal{T}$ are defined in Eqs. (4.43–4.47, page 102). It has been shown that the normal parameter set is very well adapted to the FMA, which give accurate predictions (Kosowsky et al., 2002). Since here we are interested in the predictions for $\mathcal{B} = \Omega_b h^2$ and

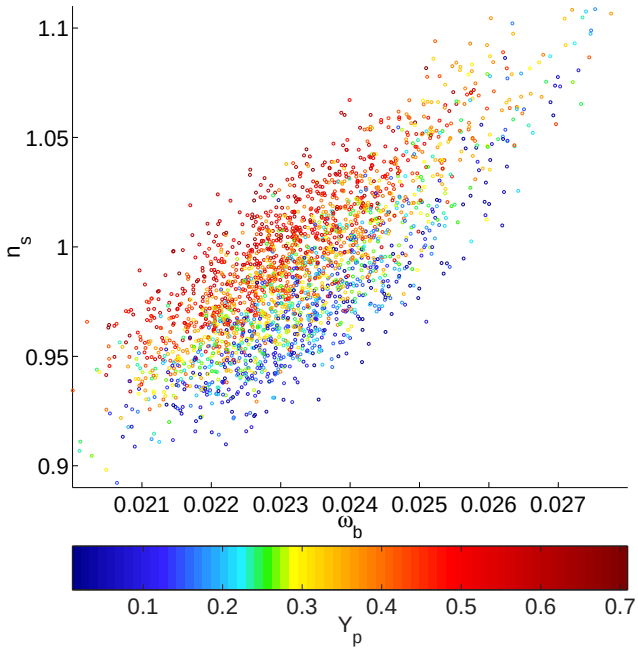

Figure 6.12: Scatter plot in the $\omega_b - n_s$ plane, with the value of Y_p rendered following the color scale. Green corresponds roughly to the BBN preferred value.

Y_p, we do not need to explicitly map the FMA forecasts in the normal parameter space onto the cosmological parameter space.

The choice of the physical parameter set makes it easy to implement in the FMA interesting theoretical priors. For instance, we are interested in imposing flatness in our forecast, in order to be able to directly compare present-day accuracy on Y_p with the potential of Planck and of and ideal CMB experiment (see below). The prior on the curvature of the universe is imposed in the FMA by fixing the value of the parameter $\mathcal{A}$ to the one of the fiducial model. In fact, the parameter $\mathcal{A}$ is a generalization of the shift parameter, which describes the sideways shift of the acoustic peak structure of the CMB power spectrum as a function of the geometry of the universe and its content in matter, radiation and cosmological constant. Although imposing $\mathcal{A} = \text{const}$ is not the same as having a constant spatial curvature over the full range of cosmological parameters, for the purpose of evaluating derivatives the two conditions reduce to the same. The fact that our fiducial model is actually slightly open (see below), does not make any substantial difference in the results, apart from reducing the numerical inaccuracies which would arise had we computed the derivatives around an exactly flat model. We can also easily impose a prior knowledge of the helium fraction, by fixing the value of Y_p, as is usually the case for present CMB analyses, and investigate how this modifies the expected error on the the baryon density.

Accuracy issues

We numerically compute double sided derivative of the power spectrum around the fiducial model with cosmological parameters given in Table 6.4. We find it necessary to increase

Parameter		Value
Baryons	Ω_b	0.046
Matter	Ω_m	0.270
Dark Energy	Ω_Λ	0.720
Radiation	$\Omega_{\rm rad}$	$7.95 \cdot 10^{-5}$
Massless ν families	N_ν	3.04
Total density	$\Omega_{\rm tot}$	0.990
Hubble constant	h	0.72
Optical depth	$\tau_{\rm re}$	0.166
Spectral index	n_s	0.99
Normalization	A_s	$2 \cdot 10^{-9}$

Table 6.4: Cosmological parameters for the fiducial ΛCDM model around which the FMA is performed. We choose a slightly open model to avoid numerical inaccuracies in the derivatives.

the accuracy of CAMB by a factor of 3 in each of the "accuracy boost" values. As a fiducial model, we use the best fit model to the WMAP data for the standard ΛCDM scenario, as given in Table 1 of Spergel et al. (2003). However, in order to avoid numerical inaccuracies which arise when differentiating around a flat model, we reduce slightly the value of Ω_Λ by imposing an open universe, $\Omega_{\rm tot} = 0.99$.

We perform the FMA for the expected capabilities of Planck's High Frequency Instrument (HFI) and for an ideal CMB measurement which would be cosmic variance limited (CVL) both in temperature and in E-polarization (and we do not consider the B-polarization spectrum), and therefore represents the best possible parameter measurement from CMB anisotropies alone. The complicated issues coming from foreground removals, point source subtraction, etc. are assumed to be already (roughly) taken into account by the experimental parameters, see § 5.2.1 for definitions. These are the effective percentage sky coverage $f_{\rm sky}$, the number of channels, the sensitivity of each channel $\sigma_c^{T,E}$ for temperature (T) and E-polarization (E) in μK and the angular resolution $\theta_c^{T,E}$ (in arcmin). For Planck HFI, we take the three channels with frequencies 100, 143 and 217 GHz, with respectively $\sigma^T_{c=1,2,3} = 5.4, 6.0, 13.1$ and $\sigma^E_{c=2,3} = 11.4, 26.7$ and we have $f_{\rm sky} = 0.85$ (Planck Website, 2004) Since the CVL is an ideal experiment, we put its noise to zero and assume perfect foregrounds removal, so that $f_{\rm sky} = 1$. In order to test the accuracy of our predictions and compare present-day results with the forecasts, we also perform an FMA with WMAP first year parameters, obtaining excellent agreement between the FMA results and the errorbars from actual data. For the purpose of comparison, we include forecasts for the full WMAP four year mission, which will also measure E-polarization and reduce present-day errors on the temperature spectrum by a factor of two. We limit the range of multipoles to $\ell < 2000$, because at smaller angular scales non-primary anisotropies begin to dominate (Sunyaev-Zeldovich effect). Seljak et al. (2003b) discuss the issue of numerical precision of three different CMB codes and conclude that they are accurate to within 0.1%. While this is encouraging, it is not of direct relevance to this work, since what matters in the computation of derivatives is not much the absolute precision of the spectra, but rather their relative accuracy.

Temperature + TE-cross + E-polarization					
	No priors		Flatness		Flatness and $Y_p = 0.24$
	$\frac{\Delta Y_p}{Y_p}$	$\frac{\Delta\omega_b}{\omega_b}$	$\frac{\Delta Y_p}{Y_p}$	$\frac{\Delta\omega_b}{\omega_b}$	$\frac{\Delta\omega_b}{\omega_b}$
WMAP 4yrs [1]	~ 50	2.92	~ 40	2.86	2.86
Planck	7.60	1.31	4.96	1.26	0.70
CVL	2.59	0.34	1.52	0.32	0.13
Temperature + TE-cross					
WMAP 1st yr [2]	N/A	N/A	71.25	5.04	5.04
WMAP 4yrs [1]	~ 75	4.10	~ 60	3.94	3.94
Planck	8.91	1.74	6.60	1.63	0.74
CVL	5.18	0.55	2.84	0.55	0.19

Table 6.5: Fisher matrix forecasts and comparison with present-day results for different priors and using different combinations of temperature and polarization CMB spectra. Errors are in percent with respect to the values of the fiducial model, $Y_p = 0.24$ and $\omega_b = 0.0238$ (1σ l.c. all other marginalized).

Forecasts and discussion

Table 6.5 summarizes our forecasts for the future measurements and compares them with the results obtained from WMAP actual data.

We notice that when the WMAP full four year data will be available (including E-polarization), the error on the baryon density is expected to decrease by a factor of two to about 3%, compared to today's 5% (assuming flatness). Nevertheless, inclusion of Y_p as a free parameter will still have no effect on the determination of ω_b for WMAP, i.e. Y_p will remain an essentially flat direction when marginalized over. While the determination of the helium fraction will improve, the FMA cannot reliably assess quantitatively how much, since for such large errors the likelihood distribution is not Gaussian and the quadratic approximation breaks down. In the table we therefore give the FMA estimation as an indication, with the caveat that the Fisher approximation is likely to be inaccurate for the real errors on Y_p from WMAP's four year data-set.

It is interesting that for Planck, the effect of the helium fraction can no longer be neglected. Inclusion of the helium fraction increases the error on ω_b by roughly 80%, from 0.7% to 1.3%. The correlation between the two parameters will have to be taken into account, as is evident from Fig. 6.13. The expected correlation coefficient is $\mathrm{corr}(Y_p, \omega_b) = 0.84 \quad (0.91)$ for Planck (for CVL). The expected 1σ error on Y_p is about 5% for Planck, or $\Delta Y_p \sim 0.01$. This is of the same order as the spread in current astrophysical measurements. We conclude that in Planck-accuracy data analysis, it will be necessary to include the uncertainty in the determination of the helium mass fraction, at least in the form of a Gaussian prior over Y_p of the type we used in the CMB data analysis presented above.

Finally, measuring CMB temperature and polarization with cosmic variance accuracy would allow Y_p to be constrained to within 1.5%, or $\Delta Y_p \sim 0.0036$ (assuming flatness). Such

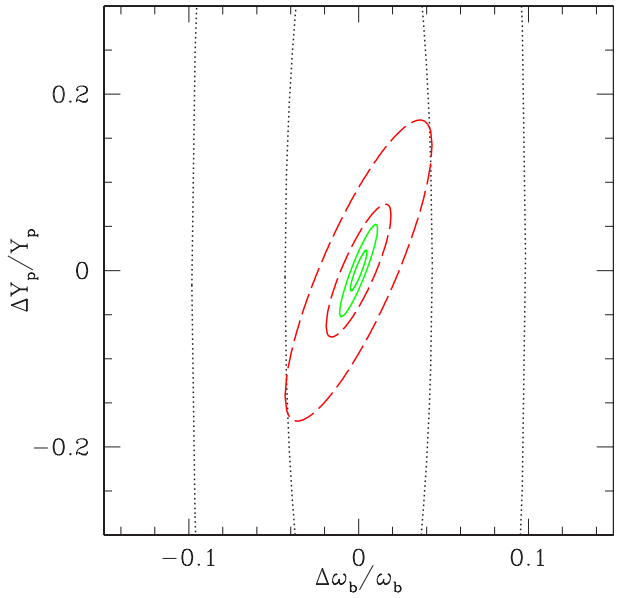

Figure 6.13: FMA forecast for the expected errors from WMAP four year mission (dotted-black), Planck (dashed-red) and a CVL experiment (solid-green). The ellipses encompass 1σ and 3σ joint likelihood regions for $\omega_b - Y_p$ (all other parameters marginalized). The axis values give the error in with respect to the fiducial model values. This forecast is for the full CMB information (Temperature, TE-cross, E-polarization) and assumes flatness.

an ideal measurement would be able to discriminate between the BBN-guided, deuterium based helium value and the current lowest, direct helium observations (cf. Fig. 6.8).

Our forecasts for the uncertainty in the Helium mass fraction from future observations are in excellent agreement with the findings of Kaplinghat et al. (2003b). There, the standard deviation on Y_p for Planck is estimated to be $\Delta Y_p = 0.012$. Kaplinghat et al. (2003b) also consider an experiment (CMBPol) with characteristics similar to our CVL, for which they forecast $\Delta Y_p = 0.0039$, again in close agreement with our result. In an earlier work, Eisenstein et al. (1998a) found for Planck (temperature and polarization) $\Delta Y_p = 0.013$, also in satisfactory concordance with our result. It should be noticed that the forecast reported for MAP in Table 2 of Eisenstein et al. (1998a), namely $\Delta Y_p = 0.02$, is nothing but the Gaussian prior $Y_p = 0.24 \pm 0.02$ which was assumed in their analysis.

The main source of improvement for the determination of Y_p will be the better sampling of the temperature damping tail provided by Planck and the CVL. Polarization measurements have mainly the effect of reducing the errors on other parameters. In fact, we have checked that excluding from our FMA the $2 \leq \ell \leq 50$ region of the E-polarization and ET-correlation spectra changes the forecast precision on Y_p less than about 10-15% for Planck and less than a few percent for CVL. This supports the conclusion that the low-ℓ reionization bump is not very useful in measuring the helium abundance, because of the degeneracy with $z_{\rm re}$.

6.3 Time variations of the fine-structure constant

The search for observational evidence for time or space variations of the 'fundamental' constants that can be measured in our four-dimensional world is an extremely exciting area of current research, with several independent claims of detections in different contexts emerging in the past few years. In particular, possible time variations of the fine-structure constant can be tested with the CMB, and represent another line of investigation going beyond the standard description of cosmology. The contents of this section summarize the latest result of a rather large collaboration I have been involved with, aimed at constraining time variations of the fine-structure constant using CMB anisotropy. We thoroughly studied the issue of crucial degeneracies with other cosmological parameters and discussing what improvements can be expected with forthcoming data-sets (Martins et al., 2002, 2004; Rocha et al., 2004).

We motivate the search for time variations of the fine-structure constant in § 6.3.1, and review the current observational status of observations other than the CMB in § 6.3.2. After presenting the relevance of the fine-structure constant for CMB anisotropies in § 6.3.3 and § 6.3.4, in § 6.3.5 we provide up-to-date WMAP constraints on the value of α at the epoch of decoupling; § 6.3.6 is dedicated to a detailed Fisher matrix analysis which encompasses the standard parameters plus the fine-structure constant for the full WMAP four year data, for the Planck satellite and for a cosmic variance limited, ideal experiment.

6.3.1 Motivation

Cosmology and astrophysics play an increasingly important role as testing ground for our understanding of fundamental physics, since they provide us with extreme conditions (that one has no hope of reproducing in terrestrial laboratories) in which to carry out a plethora of tests and search for new paradigms. Perhaps the more illuminating example is that of multidimensional cosmology: currently preferred unification theories (Polchinski, 1998; Damour, 2003a) predict the existence of additional space-time dimensions, which will have a number of possibly observable consequences, including modifications in the gravitational laws on very large (or very small) scales (Will, 2001) and space-time variations of the fundamental constants of nature (Martins, 2002; Uzan, 2003).

The most promising case, and the one that has been the subject of most recent work and speculation, is that of the fine-structure constant

$$\alpha = \frac{e^2}{\hbar c} \tag{6.26}$$

where e is the electron charge, c the speed of light and $\hbar$ Planck's constant.

There have been a number of recent reports of evidence for a time variation of fundamental constants (Webb et al., 2001, 2003; Murphy et al., 2001c; Ivanchik et al., 2003), which we review below. Apart from their obvious direct impact if confirmed, they are also crucial in a different, indirect way, since they provide us with an important (and possibly even unique) opportunity to test a number of fundamental physics models, such as string theory. Indeed here the issue is not *if* such a theory predicts such variations, but *at what level* it does so, and hence if there is any hope of detecting them in the near future, or if we have done it already.

On the other hand, the theoretical expectation in the simplest, best motivated model is that α should be a non-decreasing function of time (Damour & Nordtvedt, 1993; Santiago et al., 1998; Barrow et al., 2002). This is based on rather general and simple assumptions, in particular that the cosmological dynamics of the fine-structure constant is governed by a scalar field whose behavior is akin to that of a dilaton. If this is so, then it is particularly important to try to constrain it at earlier epochs, where any variations relative to the present-day value should be larger. However, one of the interpretations of the Oklo results is that α was *larger* at an epoch corresponding to a redshift of about $z \sim 0.1$ than today, whereas the quasar results indicate that α was *smaller* at $z \sim 2-3$ than today, see below for more details. If both results are validated by future experiments, then the above theoretical expectation must clearly be wrong, which would be a perfect example of using astrophysics to learn about fundamental physics. Playing devil's advocate, one could certainly conceive that cosmological observations of this kind could one day prove string theory wrong. Indeed, it has been argued (Damour, 2003a,b) that even the results of Webb and collaborators may be hard to explain in the simplest, best motivated models where the variation of the fine-structure constant is driven by the spacetime variation of a very light scalar field.

Cosmic microwave background anisotropies provide a tool to measure the fine-structure constant at high redshift, being mostly sensitive to the epoch of decoupling, $z \sim 1100$.

6.3.2 The observational status

The recent explosion of interest in the study of varying constants is mostly due to the results of Webb and collaborators (Murphy et al., 2001b; Webb et al., 2001; Murphy et al., 2001c,a) of a 4σ detection of a fine-structure constant that was smaller in the past,

$$\frac{\Delta\alpha}{\alpha} = (-0.72 \pm 0.18) \times 10^{-5}\,, \quad z \sim 0.5 - 3.5\,; \tag{6.27}$$

indeed, more recent work (Murphy et al., 2003; Webb et al., 2003) provides an even stronger detection. These results are obtained through comparisons of various transitions (involving various different atoms) in the laboratory and in quasar absorption systems, using the fact that the size of the relativistic corrections goes as $(\alpha Z)^2$. A number of tests for possible systematic effects have been carried out, all of which have been found either not to affect the results or to make the detection even stronger if corrected for.

A somewhat analogous (though simpler) technique uses molecular hydrogen transitions in damped Lyman-α systems to measure the ratio of the proton and electron masses, $\mu = m_p/m_e$ (using the fact that electron vibro-rotational lines depend on the reduced mass of the molecule, and this dependence is different for different transitions). The latest results (Ivanchik et al., 2002) using two systems at redshifts $z \sim 2.3$ and $z \sim 3.0$ are

$$\frac{\Delta\mu}{\mu} = (5.7 \pm 3.8) \times 10^{-5}\,, \tag{6.28}$$

or

$$\frac{\Delta\mu}{\mu} = (12.5 \pm 4.5) \times 10^{-5}\,, \tag{6.29}$$

depending on which of the (two) available tables of "standard" laboratory wavelengths is used. This implies a 1.5σ detection in the more conservative case, though it also casts some

doubts on the accuracy of the laboratory results, and on the influence of systematic effects in general.

We should also mention a recent re-analysis (Fujii, 2002) of the well-known Oklo bound (Damour & Dyson, 1996). Using new Samarium samples collected deeper underground (aiming to minimize contamination), these authors again provide two possible results for both α and the analogous coupling for the strong nuclear force, α_s,

$$\frac{\dot{\alpha}}{\alpha} \sim \frac{\dot{\alpha}_s}{\alpha_s} = (0.4 \pm 0.5) \times 10^{-17}\mathrm{yr}^{-1} \tag{6.30}$$

or

$$\frac{\dot{\alpha}}{\alpha} \sim \frac{\dot{\alpha}_s}{\alpha_s} = -(4.4 \pm 0.4) \times 10^{-17}\mathrm{yr}^{-1}\,. \tag{6.31}$$

Note that these are given as rates of variation, and effectively probe timescales corresponding to a cosmological redshift of about $z \sim 0.1$. Unlike the case above, these two values correspond to two possible physical branches of the solution. See Fujii (2002) for a discussion of why this method yields two solutions (and also note that these results have opposite signs relative to previously published ones, Fujii et al., 2000). While the first of these branches provides a null result, (6.31) is a strong detection of an α that was *larger* at $z \sim 0.1$, that is a relative variation that is opposite to Webb's result (6.27). Even though there are some hints (coming from the analysis of other Gadolinium samples) that the first branch is preferred, this is by no means settled and further analysis is required to verify it.

Still we can speculate about the possibility that the second branch turns out to be the correct one. Indeed this would definitely be the most exciting possibility. While in itself this wouldn't contradict Webb's results (since Oklo probes much smaller redshift and the suggested magnitude of the variation is smaller than that suggested by the quasar data), it would have striking effects on the theoretical modelling of such variations. In fact, proof that α was once larger than today's value would sound the death knell for any theory which models the varying α through a scalar field whose behaviour is akin to that of a dilaton. Examples include Bekenstein's theory (Bekenstein, 1982) or simple variations thereof (Sandvik et al., 2002; Olive & Pospelov, 2002). Indeed, one can quite easily see (Damour & Nordtvedt, 1993; Santiago et al., 1998) that in any such model having sensible cosmological parameters and obeying other standard constraints, α must be a monotonically increasing function of time. Since these dilatonic-type models are arguably the simplest and best-motivated models for varying α from a particle physics point of view, any evidence against them would be extremely exciting, since it would point towards the presence of significantly different, yet undiscovered physical mechanisms.

Finally, we also mention that there have been recent proposals (Braxmaier et al., 2001) of more accurate laboratory tests of the time independence of α and μ using monolithic resonators, which could improve current bounds by an order of magnitude or more.

However, given that there are both theoretical and experimental reasons to expect that any recent variations will be small, it is important to develop tools allowing us to measure α in the early universe, as variations with respect to the present value could be much larger then.

6.3.3 Effects of α on the ionization history

The reason why the CMB can be used as a probe of variations of the fine-structure constant is that these alter the ionization history of the universe. Here we present the dominant effects, see Hannestad (1999); Kaplinghat et al. (1999) for a detailed treatment.

The impact of the fine-structure constant on the CMB comes from the dependence of the differential optical depth $\dot{\tau}$ (2.117, page 52) on the Thomson scattering cross section, which is

$$\sigma_T = \frac{8\pi\alpha^2\hbar^2}{3m_e^2c^2}\,, \tag{6.32}$$

where we have reintroduced the speed of light c and the Planck constant $\hbar$, and m_e is the electron mass. Now the equilibrium electron ionization fraction $x_e^{\rm eq} \equiv n_e/n_H$ goes approximately as

$$x_e^{\rm eq} \propto \left(\frac{m_e}{T}\right)^{3/2} \exp(-B/T)\,, \tag{6.33}$$

where B is the Hydrogen binding energy

$$B = \alpha^2 m_e c^2/2 \tag{6.34}$$

(see e.g. Kolb & Turner, 1990). If we ignore the fact that $x_e(z)$ does not precisely track its equilibrium value, and since the exponential factor dominates near recombination, we would simply expect from $T \propto 1/a \propto z$ that the reionization fraction be just a function of z/α^2. This turns out to be approximately correct, even if the effect of the factor $(m_e/T)^{3/2}$ and the departure of x_e from $x_e^{\rm eq}$ need to be taken into account for a more precise estimation (Kaplinghat et al., 1999).

In general, around the decoupling epoch relevant for the CMB, the fine-structure constant can be expected to evolve with redshift, $\alpha = \alpha(z)$, but we can take a constant value $\alpha_{\rm dec} \equiv \alpha(z_{\rm dec})$ instead and consider it as an *effective* value averaged over the recombination process. Summarizing, there are two important changes in the reionization history brought about by a change in $\alpha_{\rm dec}$, the value of α at the recombination epoch, which are best discussed in terms of changes on the visibility function $g(z)$, defined in Eq. (3.126, page 79). A larger value of $\alpha_{\rm dec}$ with respect to α_0, its value today, implies:

- an increased redshift of last scattering: as estimated above, this follows from rescaling the reionization fractions as $z/\alpha_{\rm dec}^2$, hence decoupling happens earlier for a larger $\alpha_{\rm dec}$, which means that the sound horizon $r_s(z_{\rm dec})$, see Eq. (4.24, page 94), is smaller. As a consequence, we expect a shift of the peaks' structure to larger ℓ values, according to (4.17, page 93). This effect will be degenerate with the shift parameter $\mathcal{R}^{\rm shift}$ (4.22, page 94) or equivalently with the normal parameter $\mathcal{A}$, Eq. (4.43, page 100), as shown in Fig. 6.14. There will also be a boost of the first acoustic peak due to the increased early ISW effect, see § 4.1.2.3.

- A narrower peak of the visibility function: by increasing $\alpha_{\rm dec}$ the peak of the visibility function is moved to a larger redshift, when the expansion rate is faster

  $$\dot{T} \propto -\mathcal{H} \propto -(1+z) \tag{6.35}$$

  and thus the temperature and therefore x_e drop faster, which makes $g(z)$ narrower, see Fig. 6.15. This leads to a smaller damping scale, cf. Eq. (4.34, page 98), hence the small-scale power of the CMB spectrum increases for $\alpha_{\rm dec}/\alpha_0 > 1$.

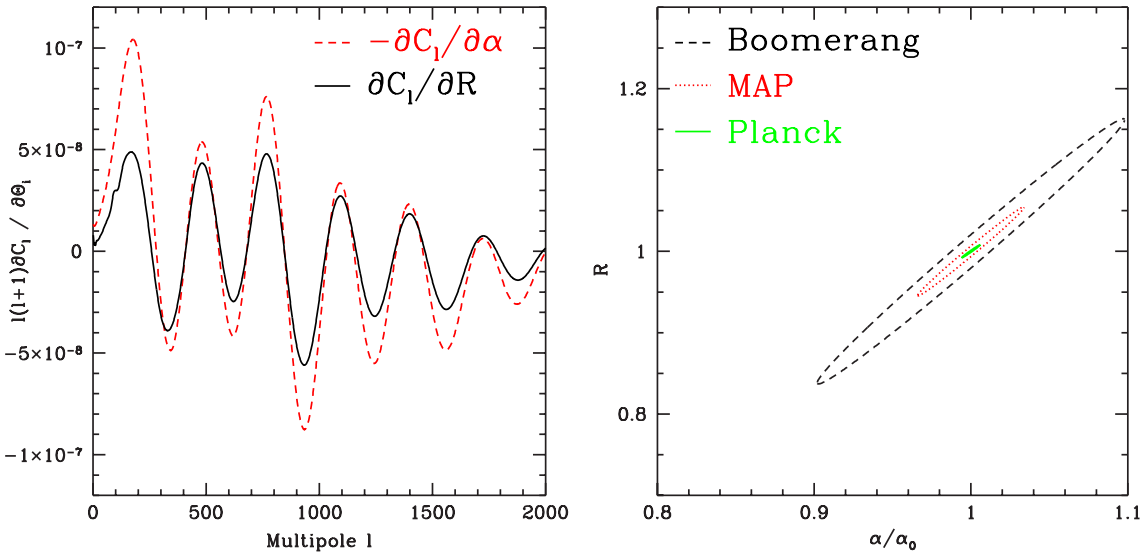

Figure 6.14: Left panel: derivatives of the temperature spectrum with respect to $\alpha_{\rm dec}$ and the shift parameter $\mathcal{R}^{\rm shift}$. We plot $-\partial C_\ell/\partial\alpha_{\rm dec}$ to facilitate the comparison with $\partial C_\ell/\partial\mathcal{R}^{\rm shift}$. The two derivatives are perfectly in phase: this is responsible for the degeneracy between the corresponding parameters (right panel, Fisher matrix analysis). Only the different amplitudes allow an experiment which maps sufficiently high multipoles with high accuracy to distinguish between them, in particular revealing the change in the damping scale brought about by changes in $\alpha_{\rm dec}$. In the right panel, the Fisher matrix results contain 1σ of the likelihood (including temperature only), and clearly indicate a strong correlation between the two parameters (see Martins et al., 2002).

In Fig. 6.16 we plot the resulting CMB temperature spectrum, where the above mentioned changes are readily distinguishable.

6.3.4 The role of reionization

After decoupling, the CMB is essentially insensitive to how α varies, until the reionization epoch is reached, at which point Thomson scattering becomes effective again. If the value of α at reionization, $\alpha_{\rm re} \equiv \alpha(z_{\rm re})$, is different from its value today, it will affect the CMB spectrum through a change in the reionization optical depth $\tau_{\rm re}$. However, $\tau_{\rm re}$ is itself dependent on the cosmological model and possibly on a number of relevant non-linear physical processes related to the astrophysical mechanisms responsible for the reionization. In general, this problem is solved by treating $\tau_{\rm re}$ as a free parameter, which accounts for the relatively poor knowledge of the details of the reionization history and in our case for the uncertainty about the exact value of α during the reionization epoch. We conclude that provided we treat $\tau_{\rm re}$ as a free parameter the lack of a precise knowledge of the value of α during the epoch of reionization is unimportant, and we can take $\alpha_{\rm re} = \alpha_0$. On the more phenomenological side, the results of Webb and collaborators for the value of α at a redshift of $2-3$ would suggest that at the epoch of reionization the possible changes in α relative to the present day are already very small. Therefore one can calculate the effect of a varying α by simply

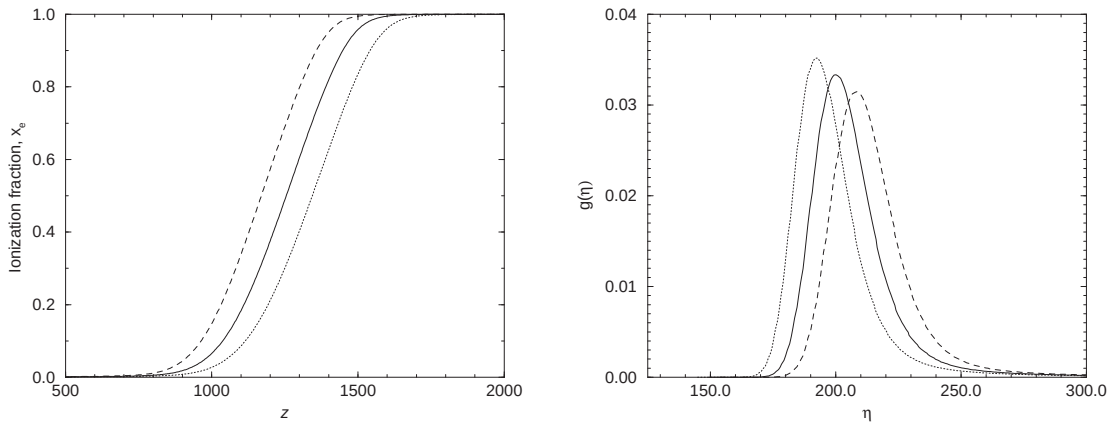

Figure 6.15: Ionization fraction as a function of redshift (left panel) and visibility function as a function of conformal time (right panel) for different values of the fine-structure constant at decoupling: $\alpha_{\rm dec}/\alpha_0 = 1$ (solid), $\alpha_{\rm dec}/\alpha_0 = 1.03$ (dotted), $\alpha_{\rm dec}/\alpha_0 = 0.97$ (dashed). Decoupling happens earlier and the last scattering surface is narrower for $\alpha_{\rm dec}/\alpha_0 > 1$.

assuming two values for the fine-structure constant, one at low redshift, $z \lesssim 20$, for which we take today's value by the above argument, and one around the epoch of decoupling, $\alpha_{\rm dec}$, which we want to determine.

As shown in § 4.1.3.2, reionization changes the amplitude of the acoustic peaks in the temperature spectrum, without affecting their position and spacing, while introducing the reionization bump at low ℓ in the polarization spectrum. If the value of $\alpha_{\rm dec}$ is different from the value today (which corresponds to $\alpha_{\rm re}$), then the peaks in the polarization power spectrum at small angular scales will be shifted sideways, while the reionization bump on large angular scales will remain fixed. This is illustrated in Fig. 6.16 (lower right panel). It follows that by measuring the separation between the acoustic peaks and the bump, one could in principle measure both α and the reionization optical depth $\tau_{\rm re}$, as shown in Fig. 6.17. This holds true as long as one assumes a specific reionization history, such as the sudden reionization scenario used here. However, if we would allow for a more realistic reionization modelling, the detailed dependence of the reionization bump on the new reionization parameters is likely to wash out this effect. Nevertheless, with present-day accuracy the CMB data are sensitive only to the optical depth of reionization, as pointed out in § 6.2.2, which justify the use of the simplest reionization modelling. Within this framework, the fact that $\tau_{\rm re}$ unexpectedly turned out to be as large as 0.16 as derived from the WMAP data (Spergel et al., 2003) makes the prospects of constraining α with the CMB much better because of the above effect.

Finally, we point out that the modifications discussed above are direct consequences of an α variation, and that indirect effects are usually present as well since any variation of α is necessarily coupled with the dynamics of the Universe (Mota & Barrow, 2004). Here we take a pragmatic approach and say that, since the CMB is insensitive to the details of α variations from decoupling to the present day, we do not in fact need to specify a redshift dependence for this variation – although we could have specified one if we so chose. At this stage, we prefer to focus on model-independent constraints, and hence do not attempt to include

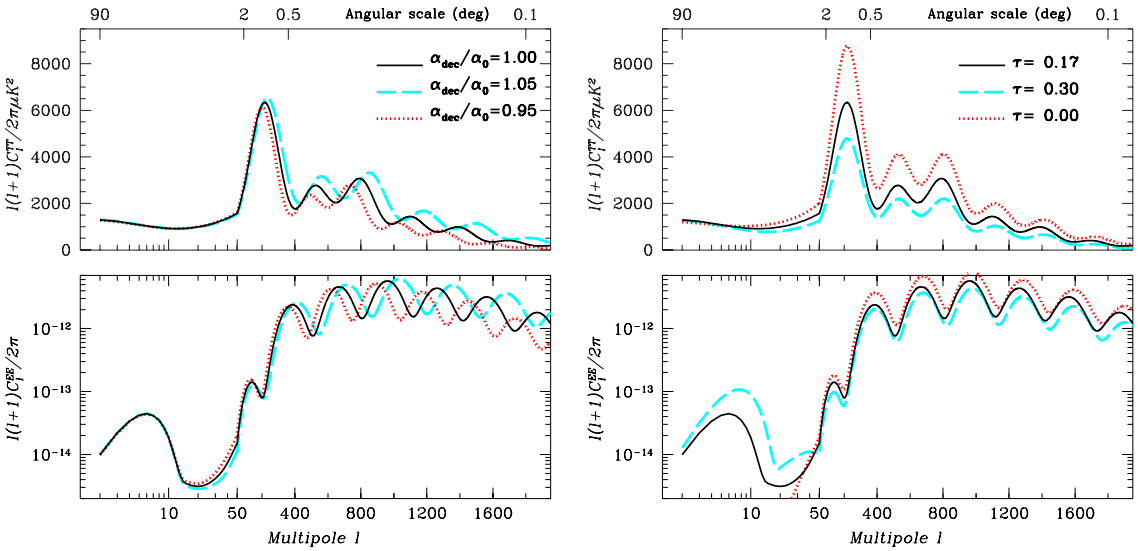

Figure 6.16: Contrasting the effects of varying $\alpha_{\rm dec}$ (left) and reionization optical depth $\tau_{\rm re}$ (right) on the CMB temperature (top) and polarization (bottom). The reionization bump is not changed by variations of $\alpha_{\rm dec}/\alpha_0$. The black lines are for the WMAP best fit model, with $\alpha_{\rm dec}/\alpha_0 = 1$ and $\tau_{\rm re} = 0.17$.

an explicit modelling for the redshift dependence $\alpha(z)$. Nevertheless, given some model-independent constraints one can always translate them into constraints on the parameters of one's favorite model. Beside possible time variations of α, investigated here, one could also envisage searching for spatial variations on the last scattering surface (Sigurdson et al., 2003).

6.3.5 CMB constraints on α from WMAP alone

We use a modified version of CMBFAST which includes the effects of varying α described above, to analyse the recent WMAP temperature and cross-polarization data adopting the likelihood estimator method described in Verde et al. (2003). The models are sampled on an uniform grid in a 7 dimensional parameter space as follows:

$$\begin{aligned}
0.05 < \Omega_c h^2 &< 0.20 \quad (0.01)\,,\\
0.010 < \Omega_b h^2 &< 0.028 \quad (0.001)\,,\\
0.500 < \Omega_\Lambda &< 0.950 \quad (0.025)\,,\\
0.900 < \alpha_{\rm dec}/\alpha_0 &< 1.050 \quad (0.005)\,,\\
0.06 < \tau_{\rm re} &< 0.30 \quad (0.02)\,,\\
0.880 < n_s &< 1.08 \quad (0.005)\,,\\
-0.15 < \frac{{\rm d}n_s}{{\rm d}\ln k} &< 0.05 \quad (0.01)\,.
\end{aligned} \tag{6.36}$$

The numbers between parentheses give the step size along each direction; n_s is the scalar spectral index of the primordial power spectrum, and ${\rm d}n_s/{\rm d}\ln k$ is the spectral index running,

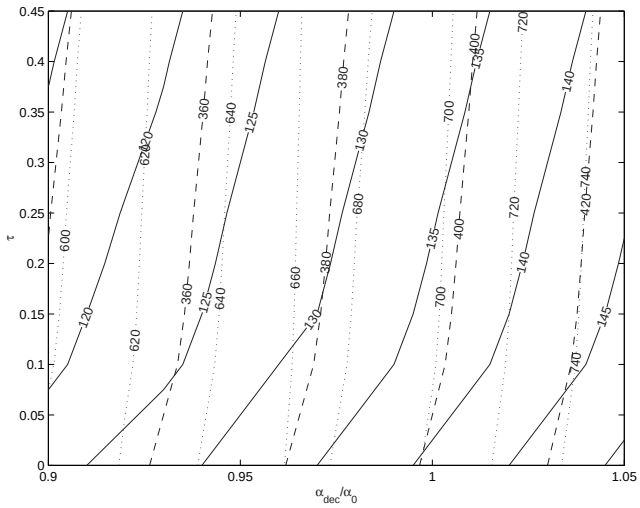

Figure 6.17: The separation in ℓ between the reionization bump and the first (solid lines), second (dashed) and third (dotted) peaks in the polarization spectrum, as a function of α at decoupling and τ. A (somewhat idealized) description of how α and $\tau_{\rm re}$ can be measured using CMB polarization.

i.e. we introduce a scale dependence of the spectral index of the form

$$n_s(k) = n_s(k_{\rm P}) + \frac{{\rm d}n_s}{{\rm d}\ln k}\ln\left(\frac{k}{k_{\rm P}}\right)\,, \tag{6.37}$$

where $n_s \equiv n_s(k_{\rm P})$ is a constant and the pivot scale $k_{\rm P}$ is chosen to be $k_{\rm P} = 0.002{\rm Mpc}^{-1}$. We only include flat models, so that the Hubble parameter $H_0 \equiv 100h$ km s^{-1} Mpc^{-1} is a derived quantity. We don't consider gravity waves or isocurvature modes since these further modifications are not required by the WMAP data.

The likelihood distribution function for $\alpha_{\rm dec}/\alpha_0$, obtained after marginalization over the remaining parameters, see § 5.1.5, is plotted in Fig. 6.18, and gives the marginalized confidence interval

$$0.95 < \alpha_{\rm dec}/\alpha_0 < 1.02 \quad \text{(at 95\% l.c.)}. \tag{6.38}$$

If we impose ${\rm d}n_s/{\rm d}\ln k = 0$ we obtain instead

$$0.94 < \alpha_{\rm dec}/\alpha_0 < 1.01 \quad \text{(at 95\% l.c.)}. \tag{6.39}$$

It is interesting to consider the correlations between a α/α_0 and the other parameters in order to see how this modification to the standard model can change our conclusions about cosmology. In Fig. 6.19 we plot the likelihood contours in the $\alpha/\alpha_0 - \tau_{\rm re}$ plane for two cases: using the temperature only WMAP data and including the TE cross correlation data. There is a clear degeneracy between these two parameters if one uses only temperature information: increasing the optical depth allows for an higher value of the spectral index n_S and a lower value of α/α_0. Inclusion of the TE data is already able to partially break this degeneracy, but, as we explain below, more detailed measurements of the polarization spectra are needed to constraint separately the two parameters,

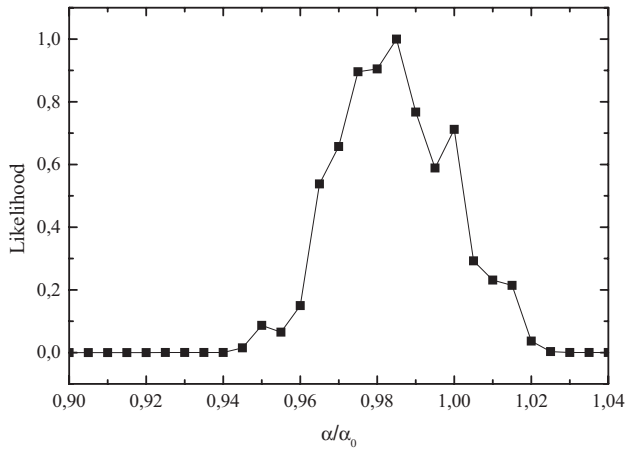

Figure 6.18: Marginalized likelihood distribution function for variations in the fine-structure constant at the time of decoupling obtained by an analysis of the WMAP data (TT+TE, one-year).

One of the most unexpected results from the WMAP data is the hint for a scale-dependence of the spectral index n_s (see e.g. Peiris et al., 2003; Kinney et al., 2004). Such a dependence should not be detectable in most of the viable single field inflationary models and, if confirmed, would have strong consequences on the possibilities of reconstructing the inflationary potential. For this reason we included the running of the spectral index in our parameter set. In Fig. 6.20 we plot likelihood contours in the $\alpha/\alpha_0 - \mathrm{d}n_s/\mathrm{d}\ln k$ plane, showing that a lower value of α/α_0 would prefer the absence of running. As already pointed out in Bean et al. (2003), a modification of the recombination scheme can therefore provide a possible explanation for the large value of $\mathrm{d}n_s/\mathrm{d}\ln k$ found from WMAP data.

In previous (pre-WMAP) work, CMB-based constraints on α were obtained with the help of additional cosmological data-sets and priors, as in Martins et al. (2002). This procedure was exposed to the criticism that different data-sets could possibly have different systematic errors that are impossible to control and could conceivably conspire to produce the results quoted. The above results are obtain from WMAP only, and therefore eliminate this possible uncertainty. For earlier works and pre-WMAP constraints, see also Avelino et al. (2000, 2001); Battye et al. (2001); Hannestad (1999).

6.3.6 Fisher matrix forecasts and degeneracies

We apply the Fisher matrix analysis (FMA) technique explained in § 5.2 to the problem of forecasting the expected precision in the determination of $\alpha_{\rm dec}$ with CMB anisotropy. For the accuracy reasons presented at length in § 5.2, § 6.1.5 and § 6.2.5, we choose to employ the following 8 dimensional base parameter set

$$\boldsymbol{\theta} = \left\{\Omega_b h^2, \Omega_m h^2, \Omega_\Lambda h^2, \mathcal{R}^{\rm shift}, n_s, Q, \tau_{\rm re}, \alpha_{\rm dec}/\alpha_0\right\} \tag{6.40}$$

which takes into account the severe geometrical degeneracy via the shift parameter $\mathcal{R}^{\rm shift}$, defined in Eq. (4.22). The quantity n_s is the scalar spectral index (without running) and Q

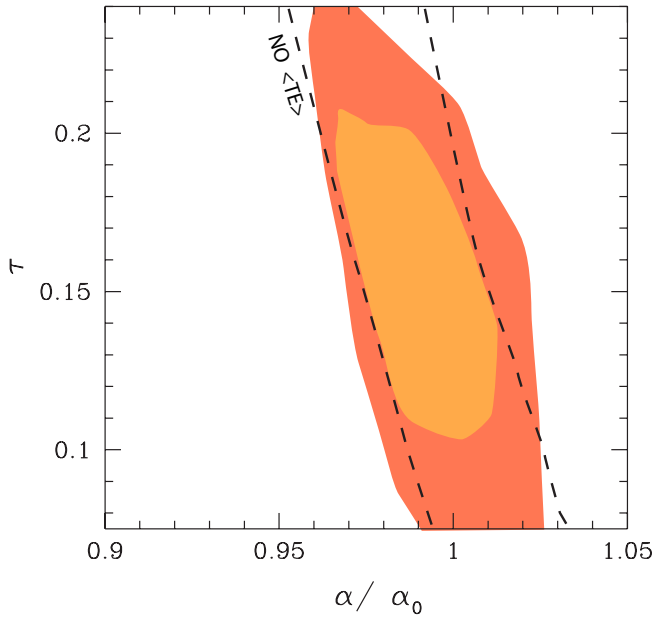

Figure 6.19: Likelihood contour plot in the $\alpha_{\rm dec}/\alpha_0 - \tau_{\rm re}$ plane including temperature information only (TT) and TT+TE together from WMAP (68% and 95% l.c. from the inside out). The inclusion of polarization data partially breaks the degeneracy between these two parameters.

a phenomenological normalization parameter as in (6.13, page 143). We restrict ourselves to scalar modes and adiabatic initial conditions.

The maximum likelihood model around which the FMA for Planck and the CVL is performed has parameters $\omega_b = 0.0200$, $\omega_m = 0.1310$, $\omega_\Lambda = 0.2957$ (and $h = 0.65$), $\mathcal{R}^{\rm shift} = 0.9815$, $n_s = 1.00$, $Q = 1.00$, $\tau = 0.20$ and $\alpha/\alpha_0 = 1.00$. We differentiate around a slightly closed model (as preferred by WMAP) with $\Omega_{\rm tot} = 1.01$ to avoid extra sources of numerical inaccuracies, since open and closed models are computed by CMBFAST using different numerical techniques which would introduce unwanted inaccuracies.

Regarding numerical accuracy issues in the computation of the Fisher matrix, we implement in the present work double–sided derivatives, which reduce the truncation error from

	WMAP			Planck		
ν (GHz)	40	60	90	100	143	217
θ_c (arcmin)	31.8	21.0	13.8	10.7	8.0	5.5
$\sigma_c T$ (μK)	19.8	30.0	45.6	5.4	6.0	13.1
σ_{cE} (μK)	28.02	42.43	64.56	n/a	11.4	26.7
$w_c^{-1} \cdot 10^{15}$ (K^2 ster)	33.6	33.6	33.6	0.215	0.158	0.350
ℓ_c	254	385	586	757	1012	1472
$\ell_{\rm max}$		1000			2000	
$f_{\rm sky}$		0.80			0.80	

Table 6.6: Experimental parameters for WMAP and Planck (nominal mission). Note that we express the sensitivities in μK. See § 5.2.1 for definitions.

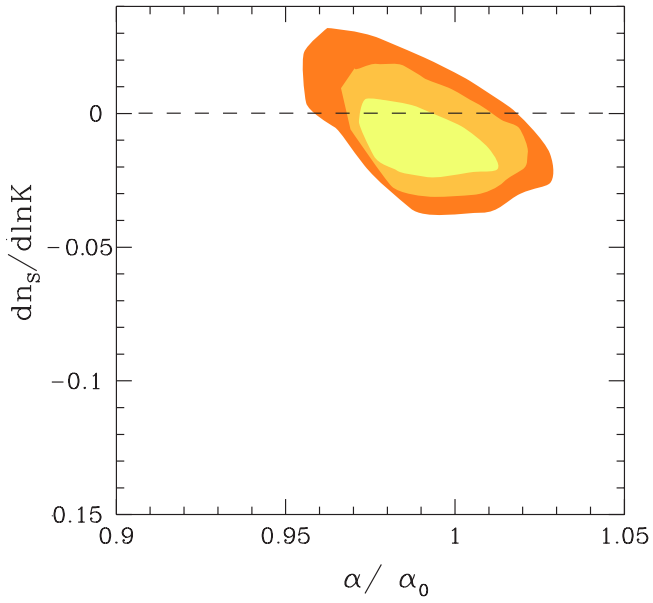

Figure 6.20: Likelihood contour plot in the $\alpha_{\rm dec}/\alpha_0 - {\rm d}n_s/{\rm d}\ln k$ plane, from WMAP temperature and ET correlation data (68%, 95% and 99% l.c. from the inside out). A zero scale dependence, as expected in most of the inflationary models, seems to be more consistent with a value of $\alpha_{\rm dec}/\alpha_0 < 1$.

second order to third order terms. The choice of the step size is a trade-off between truncation error and numerical inaccuracy dominated cases. For an estimated numerical precision of the computed models of order 10^{-4}, the step size should be approximately 5% of the parameter value (Press et al., 1992), though it turns out that for derivatives in direction of α and n_s the step size can be chosen to be as small as 0.1%. After several tests, we have chosen step sizes varying from 1% to 5% for $\omega_b, \omega_m, \omega_\Lambda$ and $\mathcal{R}^{\rm shift}$. This choice gives derivatives with an accuracy of about 0.5%. The derivatives with respect to Q are exact, being the power spectrum itself.

Predictions for WMAP's four year data

We present here the main results of the Fisher matrix forecasts; the full tables and more detailed comments can be found in Rocha et al. (2004). We first concentrate on the potential of the WMAP four year data, and we compare in Tables 6.7 and 6.8 the expected errors for two cases, for the base set of parameters (6.40) with and without inclusion of $\alpha_{\rm dec}/\alpha_0$. In both cases, we take as reference model for the Fisher matrix the WMAP best fit model of Table 1, in Spergel et al. (2003), but with a slightly larger cosmological constant which gives $\Omega_{\rm tot} = 1.01$, for the accuracy reasons explained above.

Table 6.7 gives accurate predictions for the errors on standard cosmological parameters, for models including non-flat cosmologies. Clearly, with the WMAP sensitivity, E-polarization alone will not constrain much the parameters, but combining temperature information with the polarization channels will reduce the errors on the baryon and matter density and on the shift parameter by about a factor of three, with all other parameters marginalized over. The error on the cosmological constant will remain of order unity, since

Quantity		1σ errors (%)					
		WMAP four year					
		marg.	fixed	joint	marg.	fixed	joint
		Polarization (EE)			Temperature (TT)		
baryon density	ω_b	110.64	16.58	316.44	7.33	0.81	20.96
matter density	ω_m	49.48	17.16	141.52	8.91	0.77	25.49
Λ density	ω_Λ	622.34	97.58	1779.93	113.30	83.39	324.06
spectral index	$n_{\rm s}$	69.43	4.89	198.58	6.68	0.53	19.11
normalization	Q	79.22	13.51	226.58	0.90	0.32	2.58
shift parameter	$\mathcal{R}^{\rm shift}$	46.52	13.04	133.06	9.25	0.59	26.47
reionization optical depth	$\tau_{\rm re}$	100.84	8.21	288.40	102.72	16.70	293.79
		Temp+Pol (TT+EE)			All (TT+EE+TE)		
baryon density	ω_b	2.14	0.80	6.11	2.13	0.80	6.08
matter density	ω_m	3.09	0.77	8.85	3.08	0.77	8.81
Λ density	ω_Λ	90.70	63.84	259.41	86.97	62.69	248.75
spectral index	$n_{\rm s}$	1.46	0.52	4.18	1.45	0.52	4.15
normalization	Q	0.52	0.32	1.48	0.52	0.32	1.48
shift parameter	$\mathcal{R}^{\rm shift}$	2.86	0.59	8.17	2.84	0.59	8.12
reionization optical depth	$\tau_{\rm re}$	10.52	7.45	30.08	10.41	7.44	29.78

Table 6.7: Fisher matrix analysis results for a standard model with inclusion of reionization (for the WMAP best fit model as the fisher analysis fiducial model, with $\tau_{\rm re} = 0.17$): expected 1σ errors from the WMAP-four year data. The column *marg.* gives the error with all other parameters being marginalized over; in the column *fixed* the other parameters are held fixed at their ML value; in the column *joint* all parameters are being estimated jointly.

this is an expression of the geometrical degeneracy which is fundamentally unbreakable without external priors. The spectacular improvement of about a factor 10 in determining $\tau_{\rm re}$ with polarization information is a consequence of the expected measurement of the reionization induced polarization bump, which breaks the degeneracy with normalization present with temperature alone. The spectral index accuracy thus increases by a factor 4, because the better determination of the reionization optical depth assists into breaking the small scale degeneracy with $n_{\rm s}$. The column "fixed" gives the best case scenario in which all other parameters are assumed to be known and fixed to their fiducial model value. In this case, the errors obtained by combining all channels are below 1% for all parameters but the cosmological constant.

Let us now compare this forecasts with the corresponding entries in Table 6.8, where the parameter $\alpha_{\rm dec}/\alpha_0$ has been added. The addition of a varying fine-structure constant opens up new degeneracy directions, hence the marginalized and joint error forecasts get worse (but not the errors with all other parameters fixed, of course). The most degenerate direction is with the shift parameter (marginalized errors larger by a factor 7 with all channels), as expected from the above considerations. Due to its effect on the peak heights, the fine-structure constant is largely degenerate with ω_b up to the second acoustic peak; an accurate mapping of the large multipole temperature spectrum can nevertheless lift this degeneracy,

Quantity		1σ errors (%)					
		WMAP four year					
		marg.	fixed	joint	marg.	fixed	joint
		Polarization (EE)			Temperature (TT)		
baryon density	ω_b	173.74	16.58	496.91	14.09	0.81	40.30
matter density	ω_m	260.62	17.16	745.40	13.76	0.77	39.36
Λ density	ω_Λ	637.28	97.58	1822.66	133.73	83.39	382.47
spectral index	$n_{\rm s}$	108.18	4.89	309.41	7.86	0.53	22.47
normalization	Q	96.60	13.51	276.30	2.33	0.32	6.67
shift parameter	$\mathcal{R}^{\rm shift}$	133.23	13.04	381.04	26.29	0.59	75.19
fine structure constant	$\alpha_{\rm dec}$	69.10	2.48	197.62	5.83	0.12	16.66
reionization optical depth	$\tau_{\rm re}$	228.69	8.21	654.07	103.86	16.70	297.05
		Temp+Pol (TT+EE)			All (TT+EE+TE)		
baryon density	ω_b	7.50	0.80	21.44	7.41	0.80	21.18
matter density	ω_m	5.48	0.77	15.66	5.46	0.77	15.62
Λ density	ω_Λ	91.57	63.84	261.91	87.48	62.69	250.20
spectral index	$n_{\rm s}$	2.03	0.52	5.82	2.03	0.52	5.81
normalization	Q	1.31	0.32	3.73	1.30	0.32	3.71
shift parameter	$\mathcal{R}^{\rm shift}$	14.34	0.59	41.01	14.17	0.59	40.53
fine structure constant	$\alpha_{\rm dec}$	3.08	0.11	8.80	3.05	0.11	8.71
reionization optical depth	$\tau_{\rm re}$	10.65	7.45	30.46	10.52	7.44	30.08

Table 6.8: Fisher matrix analysis results for the model of Table 6.7 with inclusion of $\alpha_{\rm dec}$.

also constraining better $n_{\rm s}$, see Martins et al. (2002) for details. This explains the larger errors on the baryon density and on the spectral index as we include α in the parameter set. However, the optical depth determination remains almost unaffected, as a consequence of the simultaneous measurement of the reionization bump's position and of the acoustic peaks angular scale, thereby validating our method for the restricted class of sudden reionization models considered here.

Predictions for Planck and an ideal experiment

We now focus on the Fisher matrix forecasts for the expected performance of the Planck satellite, and compare them with the results for an ideal CMB experiment, which would map both temperature and E-polarization with cosmic variance limited (CVL) accuracy up to $\ell = 2000$. Clearly, such a measurement is not feasible in practice, because of foreground removal and limited instrumental sensitivity, but it represents in principle the best possible parameters determination using CMB alone. The full results are tabulated in Table 6.9 and Table 6.10. In order to clarify the role of correlations between parameters, we plot in Figures 6.21 and 6.23 the 2σ joint likelihood contours for all couples of parameters for Planck, and in Figures 6.22 and 6.24 for the CVL experiment.

The first important fact is that E-polarization data alone from Planck will constrain the standard parameters better than the four year WMAP temperature data alone, compare

Table 6.7 with Table 6.9. This follows from the fact that the polarization spectrum is less plagued by large scale degeneracies than the temperature spectrum. Furthermore, as apparent from Fig. 6.21, degeneracy directions for the temperature spectrum are in many cases almost orthogonal to the directions in the polarization channel. This is especially the case for $\tau_{\rm re}$, and in fact combining temperature and polarization information reduces its marginalized error from 16% (6%) for temperature (polarization) alone to 4%. In general, the WMAP four year error-bars will be approximately halved for all parameters by Planck. Another significant aspect is that by comparing the temperature only column for Planck to the one for the CVL experiment, we conclude that Planck will be essentially cosmic variance limited as far as the temperature spectrum is concerned. This is not the case for the polarization channel, for which there will still be room for a substantial improvement over Planck's capabilities: the CVL experiment can do better than Planck by a factor 5 or more on average. The comparison of Figures 6.21 and 6.22 immediately confirms this conclusion, which makes a strong case for a post-Planck, polarization-dedicated experiment.

When we add the fine-structure constant to the Planck parameter set, the ellipses for temperature and polarization get larger for all the couples of parameters involving degenerate directions with α, compare Fig. 6.23 with Fig. 6.21. As before, this happens mostly for the $\mathcal{R}^{\rm shift}$, $n_{\rm s}$ and $\tau_{\rm re}$ using temperature information only. The degradation of the accuracy on those parameters is less dramatic than for WMAP, because Planck will map the spectrum to larger multipoles. It is remarkable that the *combined* temperature and polarization error does not grow very much when we add α, because the degeneracies are in different directions for the two channels. The fine-structure constant is the only parameter which Planck will constrain better with temperature only (0.7%) than with polarization only (2.7%, all others marginalized), while the situation is opposite for $\tau_{\rm re}$, 27% for temperature and 9% for polarization. Combining the two channels again lifts most of the degenerate directions, and we conclude that Planck will achieve an accuracy on $\alpha_{\rm dec}$ of order 0.3% (1σ, all others marginalized), thus improving by about a factor of 10 on the expected performance of the four year WMAP mission and a factor of 5 on the current upper bound (obtained however under the assumption of flatness). At the same time, the reionization optical depth will be constrained to about 4.5%. Our findings for $\alpha_{\rm dec}/\alpha_0$ and $\tau_{\rm re}$ are summarized in Fig. 6.25, where we compare degeneracy directions in the $\alpha_{\rm dec}/\alpha_0, \tau_{\rm re}$ plane for temperature alone, polarization alone and the combined channels, for Planck and the CVL experiment. We also superimpose the corresponding forecast for the WMAP four year mission (all channels) in order to facilitate the comparison.

The columns in Table 6.10 regarding the CVL experiment and the corresponding Fig. 6.24 give information about further improvements on Planck's parameter accuracy. As mentioned, a cosmic variance limited measurement of polarization could further reduce Planck's error-bars by a factor 2 to 3, reaching the highest possible accuracy from CMB alone. In particular, our analysis indicate that CMB alone can constrain variations of α up to $\mathcal{O}(10^{-3})$ at $z \sim 1100$. Going beyond will require additional priors on the other parameters.

Quantity		1σ errors (%)					
		Planck HFI			CVL		
		marg.	fixed	joint	marg.	fixed	joint
		Polarization only (EE)					
baryon density	ω_b	6.21	1.11	17.75	0.48	0.25	1.38
matter density	ω_m	3.37	0.39	9.64	0.70	0.03	1.99
cosmological constant density	ω_Λ	37.37	22.87	106.89	11.40	9.99	32.61
spectral index	n_s	1.53	0.96	4.38	0.30	0.08	0.86
normalization	Q	2.23	0.51	6.38	0.24	0.07	0.67
shift parameter	$\mathcal{R}^\mathrm{shift}$	3.33	0.35	9.52	0.65	0.03	1.86
reionization optical depth	τ_re	5.74	2.78	16.42	1.81	1.52	5.18
		Temperature only (TT)					
baryon density	ω_b	0.86	0.60	2.46	0.57	0.38	1.64
matter density	ω_m	1.51	0.13	4.31	1.10	0.08	3.14
cosmological constant density	ω_Λ	110.15	96.15	315.03	98.15	86.00	280.72
spectral index	n_s	0.54	0.13	1.56	0.36	0.07	1.04
normalization	Q	0.20	0.11	0.56	0.17	0.07	0.50
shift parameter	$\mathcal{R}^\mathrm{shift}$	1.47	0.12	4.21	1.05	0.07	3.01
reionization optical depth	τ_re	16.50	8.28	47.20	14.02	5.89	40.09
		Temperature and Polarization (TT+EE)					
baryon density	ω_b	0.80	0.53	2.30	0.32	0.21	0.92
matter density	ω_m	1.24	0.12	3.55	0.55	0.03	1.58
cosmological constant density	ω_Λ	30.58	22.04	87.46	10.72	9.85	30.65
spectral index	n_s	0.43	0.13	1.23	0.20	0.05	0.58
normalization	Q	0.19	0.10	0.53	0.14	0.05	0.41
shift parameter	$\mathcal{R}^\mathrm{shift}$	1.22	0.11	3.48	0.52	0.03	1.49
reionization optical depth	τ_re	4.04	2.65	11.56	1.73	1.48	4.96

Table 6.9: Fisher matrix analysis results including reionization ($\tau_\mathrm{re} = 0.20$): expected 1σ errors for the Planck satellite and for cosmic variance limited (CVL) experiment.

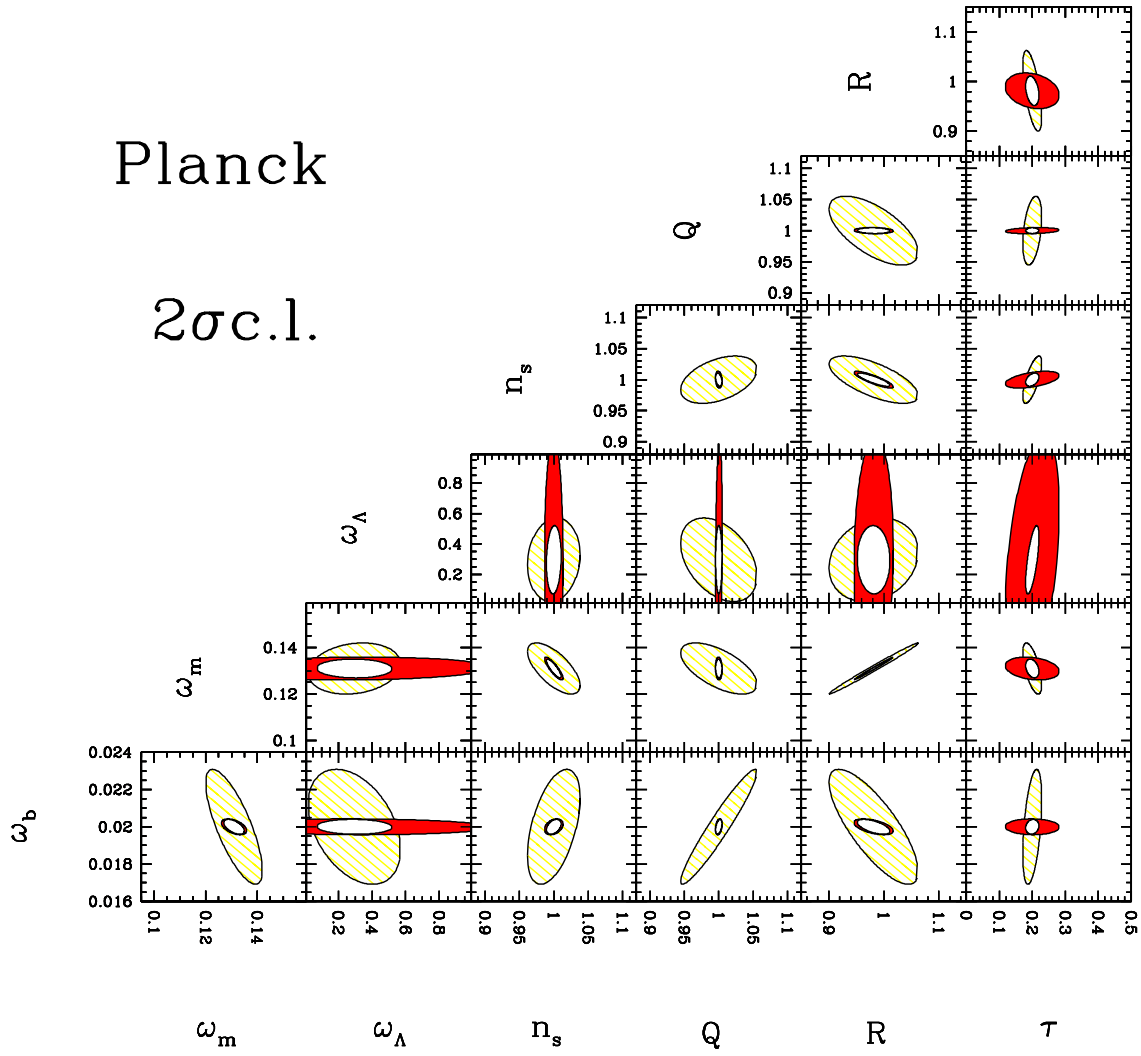

Figure 6.21: Ellipses containing 95.4% (2σ) of joint confidence (all other parameters marginalized) using temperature alone (red), E-polarization alone (yellow), and both jointly (white), for a standard model with inclusion of reionization ($\tau_{\rm re} = 0.20$). Fisher matrix forecast for the Planck HFI instrument.

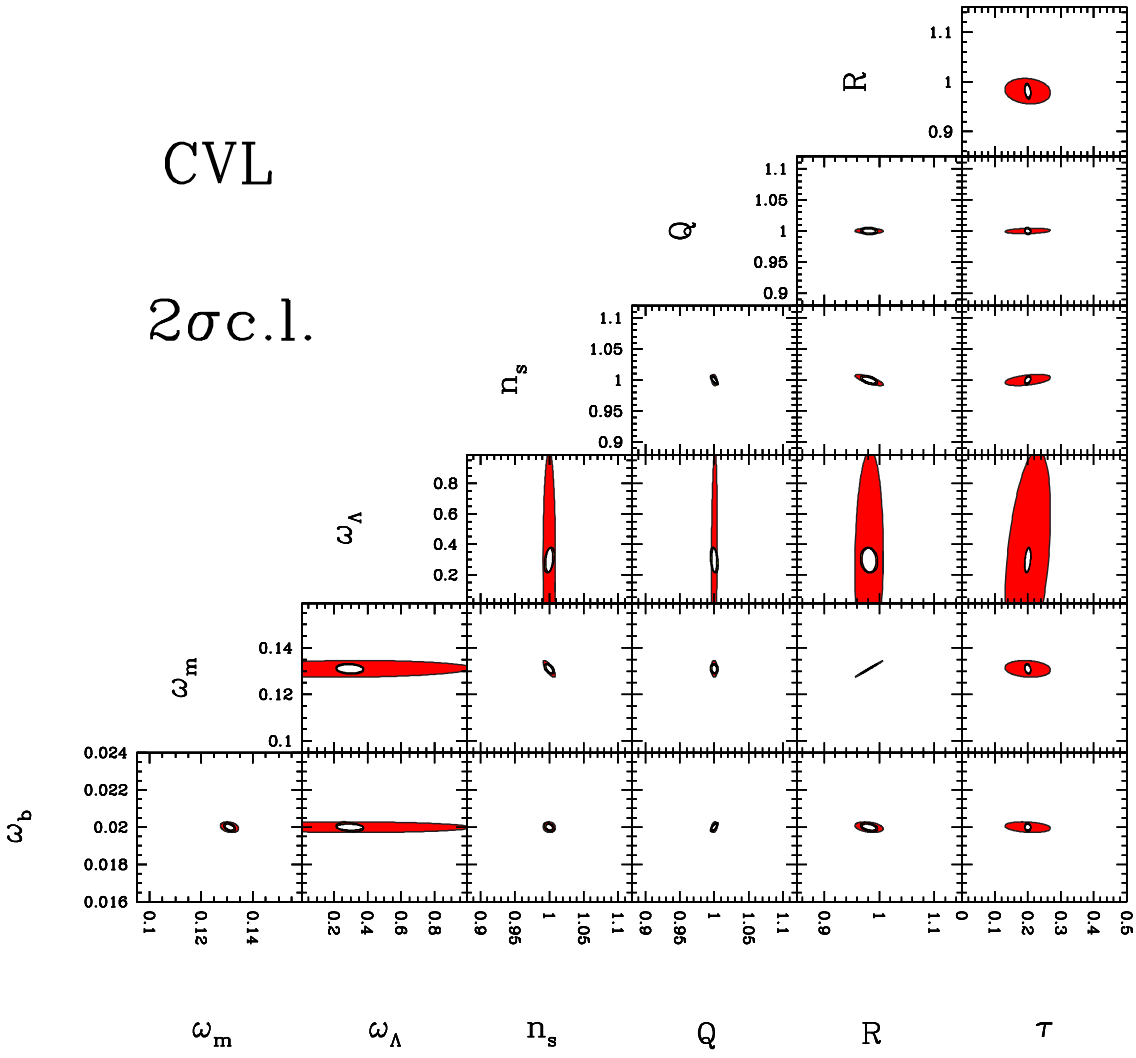

Figure 6.22: Ellipses containing 95.4% (2σ) of joint confidence (all other parameters marginalized) using temperature alone (red), E-polarization alone (yellow), and both jointly (white), for a standard model with inclusion of reionization ($\tau_{\rm re} = 0.20$). Fisher matrix forecast for an ideal cosmic variance limited (CVL) experiment.

Quantity		1σ errors (%)					
		Planck HFI			CVL		
		marg.	fixed	joint	marg.	fixed	joint
		Polarization only (EE)					
baryon density	ω_b	6.46	1.11	18.47	1.09	0.25	3.12
matter density	ω_m	7.75	0.39	22.17	1.61	0.03	4.60
cosmological constant density	ω_Λ	41.61	22.87	119.01	11.60	9.99	33.17
spectral index	n_s	4.14	0.96	11.85	0.77	0.08	2.22
normalization	Q	2.99	0.51	8.55	0.24	0.07	0.68
shift parameter	$\mathcal{R}^\mathrm{shift}$	9.56	0.35	27.33	1.19	0.03	3.40
fine structure constant	α_dec	2.66	0.06	7.62	0.40	< 0.01	1.14
reionization optical depth	τ_re	8.81	2.78	25.19	2.26	1.52	6.45
		Temperature only (TT)					
baryon density	ω_b	1.09	0.60	3.12	0.83	0.38	2.37
matter density	ω_m	3.76	0.13	10.74	2.64	0.08	7.55
cosmological constant density	ω_Λ	111.61	96.15	319.21	98.97	86.00	283.05
spectral index	n_s	2.18	0.13	6.24	1.49	0.07	4.26
normalization	Q	0.20	0.11	0.57	0.18	0.07	0.50
shift parameter	$\mathcal{R}^\mathrm{shift}$	1.58	0.12	4.53	1.06	0.07	3.04
fine structure constant	α_dec	0.66	0.02	1.88	0.41	0.01	1.18
reionization optical depth	τ_re	26.93	8.28	77.02	20.32	5.89	58.11
		Temperature and Polarization (TT+EE)					
baryon density	ω_b	0.91	0.53	2.61	0.38	0.21	1.09
matter density	ω_m	1.81	0.12	5.17	0.67	0.03	1.91
cosmological constant density	ω_Λ	30.89	22.04	88.36	10.79	9.85	30.85
spectral index	n_s	0.97	0.13	2.77	0.33	0.05	0.93
normalization	Q	0.19	0.10	0.54	0.14	0.05	0.41
shift parameter	$\mathcal{R}^\mathrm{shift}$	1.43	0.11	4.08	0.60	0.03	1.72
fine structure constant	α_dec	0.34	0.02	0.97	0.11	< 0.01	0.32
reionization optical depth	τ_re	4.48	2.65	12.80	1.80	1.48	5.15

Table 6.10: Fisher matrix analysis results as in Table 6.9 but including α_dec.

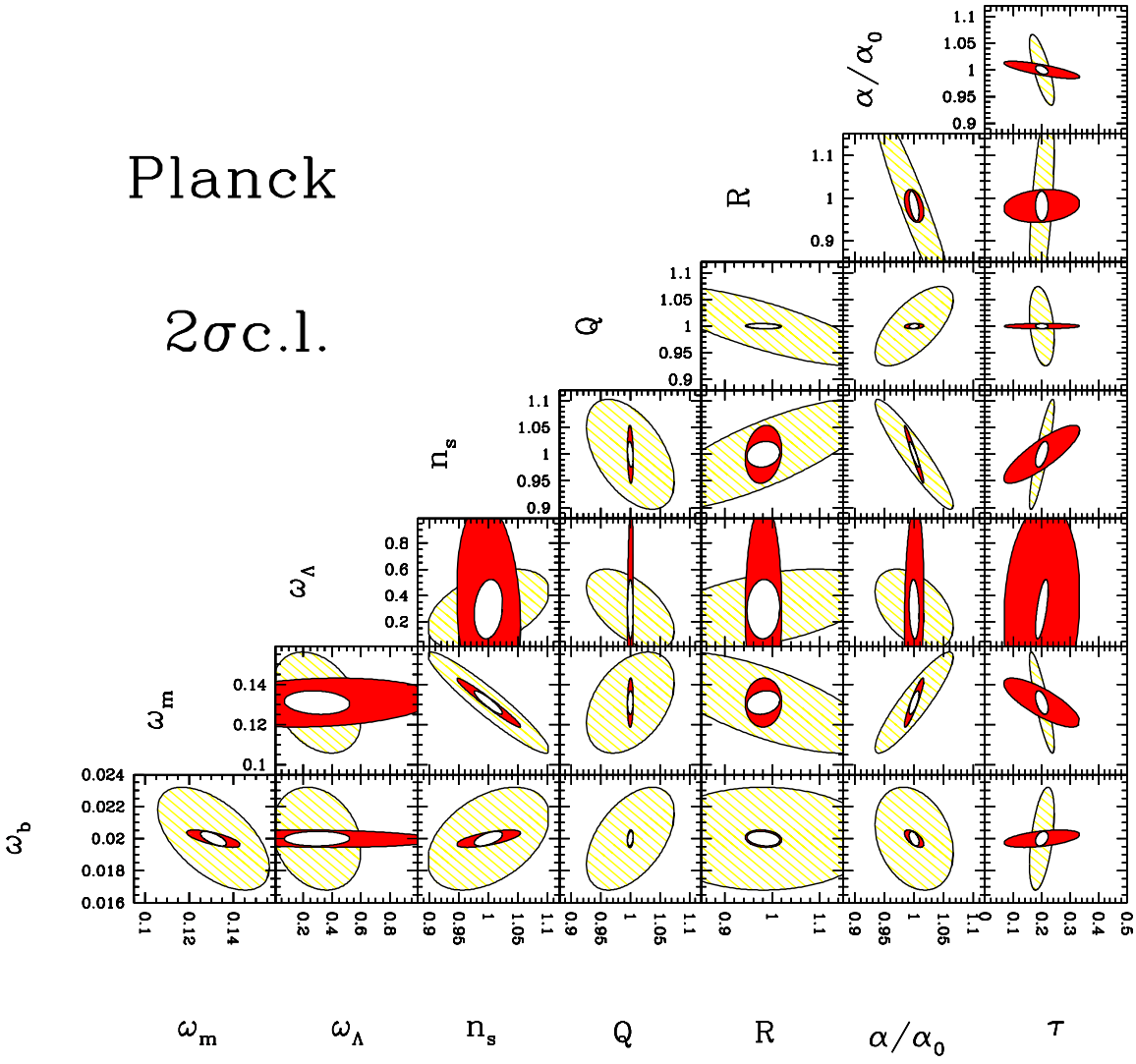

Figure 6.23: Ellipses containing 95.4% (2σ) of joint confidence (all other parameters marginalized) using temperature alone (red), E-polarization alone (yellow), and both jointly (white), for a standard model with inclusion of reionization ($\tau_{\rm re} = 0.20$) and time variations of the fine-structure constant. Fisher matrix forecast for the Planck HFI instrument.

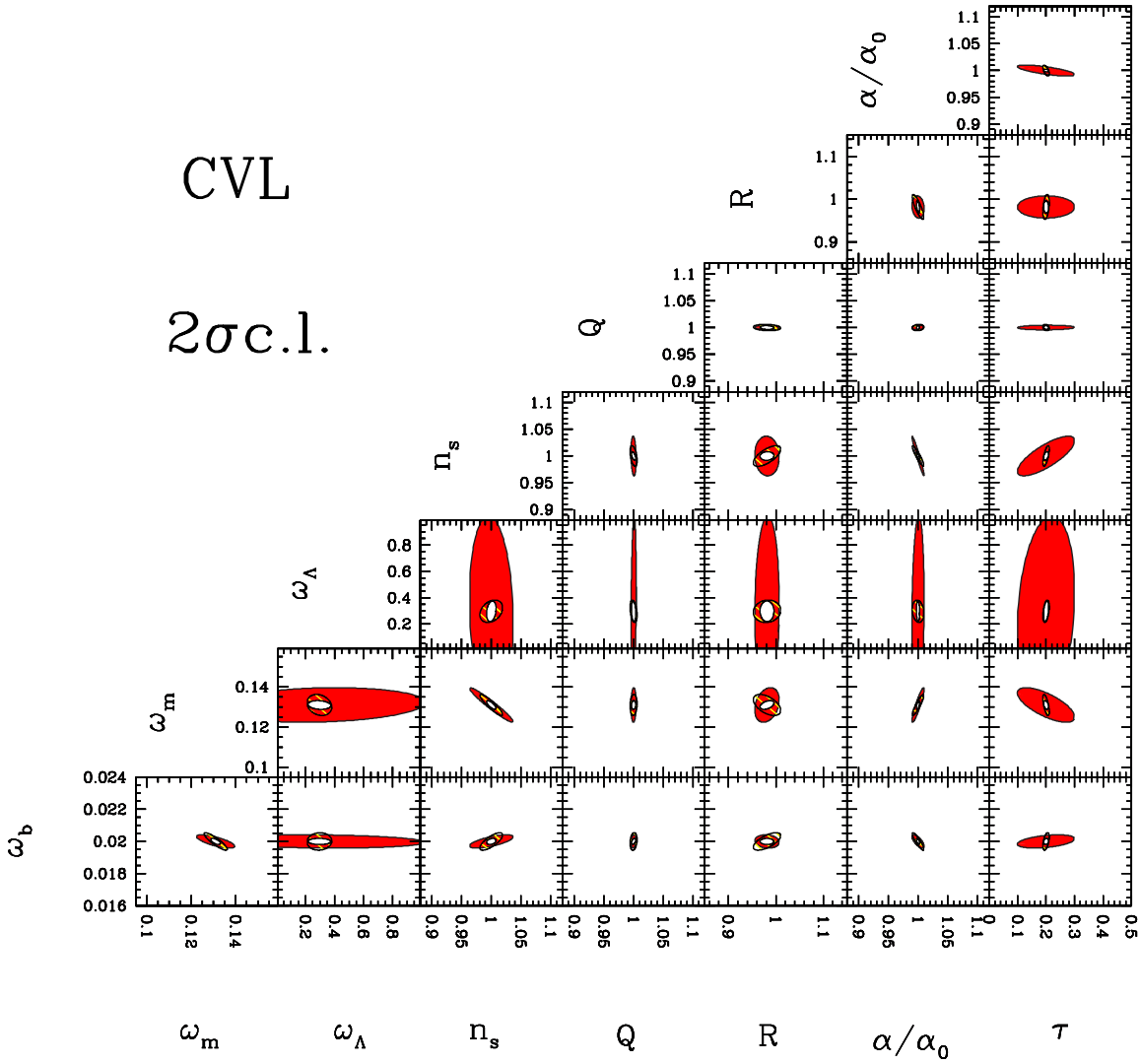

Figure 6.24: Ellipses containing 95.4% (2σ) of joint confidence (all other parameters marginalized) using temperature alone (red), E-polarization alone (yellow), and both jointly (white), for a standard model with inclusion of reionization ($\tau_{\rm re} = 0.20$) and time variations of the fine-structure constant. Fisher matrix forecast for an ideal cosmic variance limited (CVL) experiment.

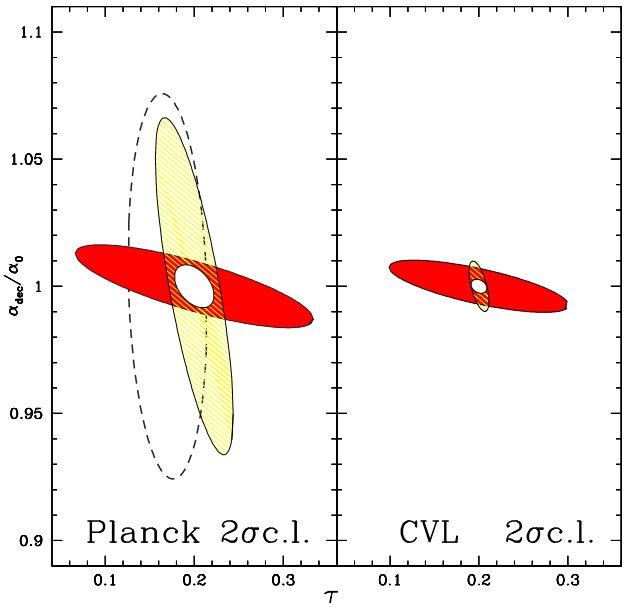

Figure 6.25: Ellipses containing 95.4% (2σ) of joint likelihood in the $\alpha_{\rm dec}/\alpha_0 - \tau_{\rm re}$ plane (all other parameters marginalized), for the Planck and cosmic variance limited (CVL) experiments, using temperature alone (red), E-polarization alone (yellow), and both jointly (white). The dashed contour represents the WMAP - 4years forecast using (TT+EE+TE) jointly.

Chapter 7

Testing the paradigm of adiabaticity

Combination of today's high quality CMB data with other cosmological data sets allows us to constrain the eight parameters

$$\boldsymbol{\theta} = \{\Omega_{\rm cdm}, \Omega_b, \Omega_\Lambda, N_\nu, h, \tau_{\rm re}, n_s, A_s\} \tag{7.1}$$

with an accuracy of a few percent (Tegmark et al., 2004b), if we assume flatness, i.e. by imposing $\Omega_\mathcal{K} = 0$. This is a spectacular achievement, even more so given the fact that many completely independent measurements seem to be converging towards the same values. In the previous sections we have discussed the determination of most of the above parameters; here we highlight that the accuracy of parameter extraction depends crucially on the assumption that the initial conditions for the perturbations are purely adiabatic, and explore the consequences of relaxing this strong assumption by including the most general type of initial conditions in the problem.

This chapter is organized as follows: we first present an introductory survey on recent CMB analysis involving isocurvature modes, § 7.1; we then investigate in a specific example how the inclusion of isocurvature modes spoils the precise determination of the baryon density from pre-WMAP CMB data in § 7.2; in § 7.3 we ask whether the presence of non-adiabatic contribution can reproduce CMB and large scale structure observations without the need for a cosmological constant, and we conclude that $\Omega_\Lambda \neq 0$ is robust with respect to the inclusion of isocurvature modes and to the use of a frequentist (rather than Bayesian) approach; finally, in § 7.4 we give the future prospects for the determination by WMAP and Planck of cosmological parameters independent of any assumption about the type of initial conditions.

7.1 Introductory survey

Until recently, most of the literature has focused on parameter extraction assuming purely adiabatic initial conditions, because the evidence for a first acoustic peak around $\ell \approx 220$ very soon ruled out the possibility of the simplest alternative, namely purely isocurvature CDM initial conditions, see e.g. Enqvist et al. (2000). Nevertheless, subdominant CDM isocurvature contributions cannot be excluded, and the constraints are even less stringent

if one allows for a correlated mixture, in which case the correlator can cancel out most of the isocurvature contribution on large scale (Langlois & Riazuelo, 2000; Amendola et al., 2002). This qualitative conclusion holds even after the more precise measurements of WMAP (Valiviita & Muhonen, 2003).

In the works of Bucher et al. (2001, 2002) the consequences for parameter extraction are examined when the most general initial conditions are allowed, with the conclusion that only a precise measurement of polarization would allow for the simultaneous reconstruction of cosmological parameters and of the initial conditions correlation matrix. The first attempt of including all the modes in a numerical parameter determination from real data is performed in Trotta et al. (2001), as illustrated in § 7.2, with the result that the pre-WMAP CMB data can not constrain to any extent the value of the baryon density and the Hubble parameter in the general initial conditions case. After the release of the WMAP first-year data, two groups have re-investigated the question of the most general initial conditions in the wake of the improved measurements: Crotty et al. (2003a) consider a correlated mixture of the adiabatic mode with each of the isocurvature modes in turn, finding that the pre-WMAP constraints on the isocurvature contribution are significantly improved; Bucher et al. (2004) refine the analysis of Trotta et al. (2001) by using Monte Carlo methods, and simultaneously including all the isocurvature modes and six cosmological parameters, but the conclusions remained qualitatively the same. The bottom line is that the relaxing the assumption of adiabaticity spoils our ability to do precision cosmology.

The phenomenological approach gives useful hints on the "stiffness" of current data, and indeed the possibility of accommodating isocurvature modes has been considerably reduced by WMAP. Although independent of any model for the generation of perturbations, this approach has the disadvantage of introducing many new free parameters in the description of the power spectrum. To reduce this number somewhat, all analyses so far have assumed the same spectral index for all modes, an assumption which is not really motivated. Since the current CMB data are in excellent agreement with purely adiabatic initial conditions, it is not surprising however that there is no statistical evidence that such extra parameters should be non-zero. Occam's razor would therefore dictate to stick to the simplest adiabatic description, lacking any evidence for a more complicated model. However, there is no compelling reason why the physics of the early universe should boil down to only one degree of freedom.

A second reason why model-independent constraints should be regarded with care is that in any specific implementation, some of the parameters will be correlated. For instance, in the curvaton scenario (Moroi & Takahashi, 2001; Lyth & Wands, 2002; Enqvist & Sloth, 2002; Lyth et al., 2003), the adiabatic and residual isocurvature modes are always totally correlated or anti-correlated. Therefore, not only the number of extra degrees of freedom is reduced, but possibly the parameter space of the model is a highly constrained subspace of the model-independent parameter space. For this reason it is interesting to derive model-specific constraints, which are more stringent than those obtained with a general phenomenological parametrization. For instance, WMAP constraints for the curvaton model have been derived for the case of CDM and baryons isocurvature fluctuations (Gordon & Lewis, 2003; Lyth & Wands, 2003). The neutrino density mode can be generated from perturbations of the neutrino chemical potential (Lyth et al., 2003), and bounds have recently been derived for this case (Gordon & Malik, 2004). It seems more difficult to produce a neutrino velocity

mode: a working model is at present still lacking.

Despite these difficulties, the CMB represents the most promising data set to learn about the type of initial conditions realized in the observed Universe: it is our window to the very early universe.

7.2 Precision cosmology and general initial conditions

In this section, based on the work published in Trotta et al. (2001), we investigate the extent to which the determination of cosmological parameters depends on the assumptions about initial conditions. We show in a specific example how the allowed parameter range is enlarged when the usual requirement for purely adiabatic initial conditions is relaxed. In order to limit the computational effort, we have chosen to vary some cosmological parameters and keep the others fixed. We consider flat models only, and we fix the total density parameter, the total matter density and the cosmological constant density parameter as follows:

$$\begin{aligned} \Omega_{\rm tot} &\equiv \Omega_\Lambda + \Omega_{\rm m} = 1\,, \\ \Omega_{\rm m} &\equiv \Omega_{\rm cdm} + \Omega_{\rm b} = 0.3\,, \\ \Omega_\Lambda &= 0.7\,, \end{aligned} \tag{7.2}$$

where $\Omega_{\rm cdm}$ and $\Omega_{\rm b}$ are the density parameters of cold dark matter (CDM) and baryons respectively, and Ω_Λ denotes the density parameter due to a cosmological constant, $\Omega_\Lambda \equiv \Lambda/3H_0^2$, and $H_0 \equiv 100h$ km s^{-1} Mpc^{-1} is the Hubble parameter today. With Ω_Λ fixed to the above values, we then vary the Hubble parameter h, the baryon density $\omega_{\rm b} \equiv \Omega_{\rm b}h^2$ and the correlation matrix $\mathbf{M}$ which describes the most general (i.e. mixed adiabatic and isocurvature) initial conditions, as explained in § 4.3. We also fix to unity the scalar spectral index, $n_{\rm s} = 1$ for all modes and cross-correlators. Even by varying only two cosmological parameters, our parameters space is still 12-dimensional, since the initial condition correlation matrix introduces ten free amplitudes.

We also investigate the following question: what is the preferred isocurvature contribution to the perturbations? We shall see that, with pre-WMAP CMB data, this question cannot be answered without strong assumptions about the cosmological parameters.

7.2.1 Pre-WMAP data analysis

Our analysis uses the COBE (Tegmark & Hamilton, 1997) and BOOMERanG (Netterfield et al., 2002) data. For the latter, we take into account the calibration and the beam size uncertainties which treated just like two additional (normally distributed) parameters of the problem ("nuisance parameters"). The two cosmological parameters $h, \omega_{\rm b}$ are sampled on a uniform grid as follows (the number in parenthesis is the step size):

$$0.50 < h < 0.80 \quad (0.05)\,, \tag{7.3}$$

$$0.015 < \omega_{\rm b} < 0.085 \quad (0.005)\,. \tag{7.4}$$

For each grid point, we search the initial condition space by minimizing the chi-square, as explained in § 5.1.5. We look for the best fit point by using a downhill simplex method (Press et al., 1992) initiated after choosing a starting point randomly. The positive semi-definiteness

of the correlation matrix **M** is ensured by penalty functions which guarantee that the conditions (4.52, page 107) are satisfied (more details are given in Trotta, 2001). The best fit is then estimated after 15,000 minimization runs using this procedure. It turns out that the topology of the χ^2 surface on our 14-dimensional parameter space (including the two above nuisance parameters) is quite complicated with many local minima and large degeneracies, which considerably complicates the numerical search. We assume that the likelihood function is Gaussian, and we maximize instead of marginalize over the parameter we are not interested in, see § 5.1.5.

In Fig. 7.1 we show the best-fit spectra for two different choices of the cosmological parameters ω_b and h. Both of them are good fits if we allow for mixed initial conditions. On the plot we have also indicated the reduced χ^2, i.e. the value of χ^2/F, where F is the number of degrees of freedom of the fit. For a fixed choice of ω_b, h the purely adiabatic model has only three parameters (the amplitude of the adiabatic mode, and the two nuisance parameters). With 26 data points (7 from COBE and 19 from BOOMERanG) this leads to $F_\mathrm{AD} = 26 - 3 = 23$ degrees of freedom. The mixed models have a symmetric 4×4 matrix determining the initial amplitude, leading to a total of 12 parameters and hence $F_\mathrm{MIX} = 14$ degrees of freedom. If we also vary ω_b and h, the number of degrees of freedom is lowered by two. It is not surprising that for fixed values $h = 0.65$, $\omega_\mathrm{b} = 0.02$, which are well fitted by the adiabatic model, the *reduced* χ^2 of the adiabatic model is somewhat lower than the one of the mixed model, since $F_\mathrm{MIX} < F_\mathrm{AD}$ (as an example, see top panel of Fig. 7.1). For the mixed model, the *absolute* χ^2 is always lower.

For both models we determine the likelihood functions of the cosmological parameters ω_b and h by maximizing the initial conditions correlation matrix and the nuisance parameters. The result is shown in the left panel of Fig. 7.2 where the likelihood contours in the (ω_b, h) plane are indicated for purely adiabatic and for mixed (general isocurvature) models. It is remarkable the extent to which the innermost 1σ contour opens up, once we allow for isocurvature components. Strangely, the least likely region is the upper left corner which contains the value of $\omega_\mathrm{b} = 0.019 \pm 0.02$ inferred from BBN (Burles et al., 2001) and the Hubble space telescope key project value for the Hubble parameter (Freedman et al., 2001) of $h = 0.72 \pm 0.08$. Moreover, there is absolutely no upper limit for ω_b within the range investigated here! This is explained by the fact that the strongest features of a high baryon density universe, the asymmetry between even and odd acoustic peaks and the shift of the peak position due to the change in the sound velocity, can be fully compensated by an admixture of isocurvature modes (see left panel of Fig. 7.1). A very high baryon density can therefore easily be accommodated into this framework. However, for high ω_b and low h, it is difficult to find a good fit because there is not enough power in the secondary peak region due to the early integrated Sachs-Wolfe effect boosting the first peak.

We define the isocurvature content of a mixed model as

$$\gamma \equiv \frac{M_{22} + M_{33} + M_{44}}{\mathrm{tr}M}\,, \tag{7.5}$$

where M_{11} denotes the adiabatic mode amplitude. The isocurvature content of the model shown in the left panel of Fig. 7.1 is only $\gamma = 0.12$, while for the parameter choice in the right panel one has $\gamma = 0.69$. Hence, if the cosmological parameters are close to those chosen in the left panel, we can conclude that the cosmic perturbations are predominantly adiabatic.

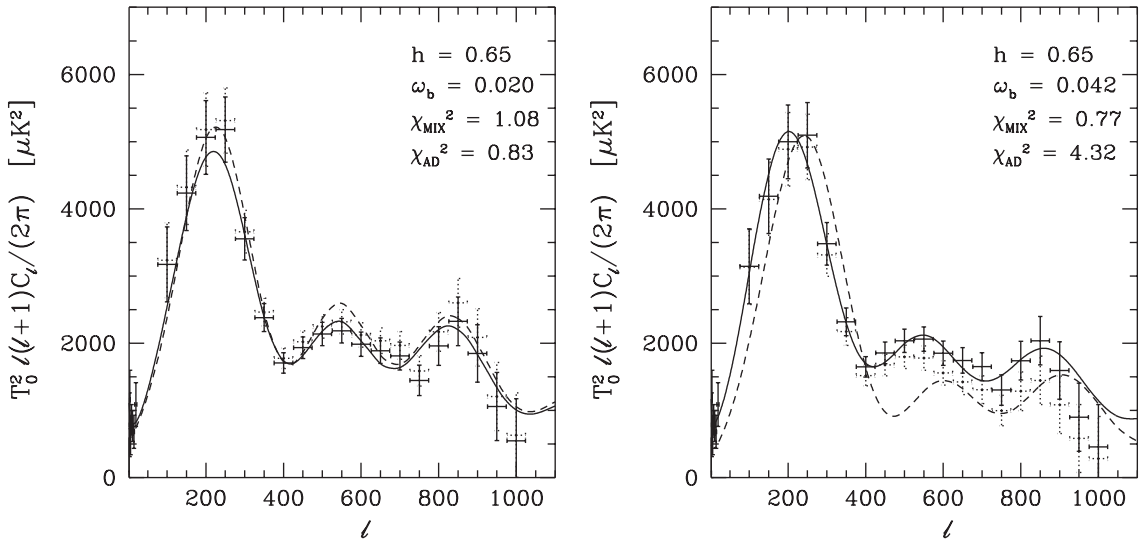

Figure 7.1: CMB anisotropy temperature spectrum for different values of the cosmological parameters $\omega_{\rm b}$ and h. We plot the best-fit corresponding to a purely adiabatic case (dashed line) and allowing general initial conditions, mixed models (solid line). The calibration and the beam size of the BOOMERanG data have been optimized to fit the mixed model (solid error bars) or the adiabatic model (dotted error bars). The parameter choice in the left panel ($\omega_{\rm b} = 0.02$, $h = 0.65$) can be fitted by both models while the values $\omega_{\rm b} = 0.042$, $h = 0.65$ (right panel), can only be fitted by a mixed model.

In the right panel of Fig. 7.2 we plot the isocurvature content, γ, of the best fit model obtained by minimizing χ^2 by variation of the initial conditions for given values of the cosmological parameters. Clearly, the further away we move from the region of parameter space well fitted by the purely adiabatic model, the higher the isocurvature contribution needed to fit the data becomes.

The main non-adiabatic component of our best fits is the neutrino entropy mode. This was to be expected, since this mode and its correlator with the adiabatic mode can shift the peak positions and can substantially add or subtract power from the second peak (Bucher et al., 2000). A crucial point is, therefore, to know whether such a mode can appear in a realistic structure formation scenario. It is known that for interacting species the non-adiabatic part of the perturbations tends to decay with time. Therefore, the generation of a neutrino entropy component can only occur after neutrino decoupling, that is at $T \lesssim 1$ MeV (see Gordon & Malik, 2004 for a discussion). A neutrino isocurvature perturbation could also be due to a fourth species of sterile neutrinos which may have decoupled very early in the history of the Universe. The same remark also applies of course to the CDM isocurvature mode. Note that the energy density of this fourth neutrino type cannot be very high, in order not to contradict the light element abundances, but there is nothing which prevents (at least in principle) the presence of large perturbations in this component.

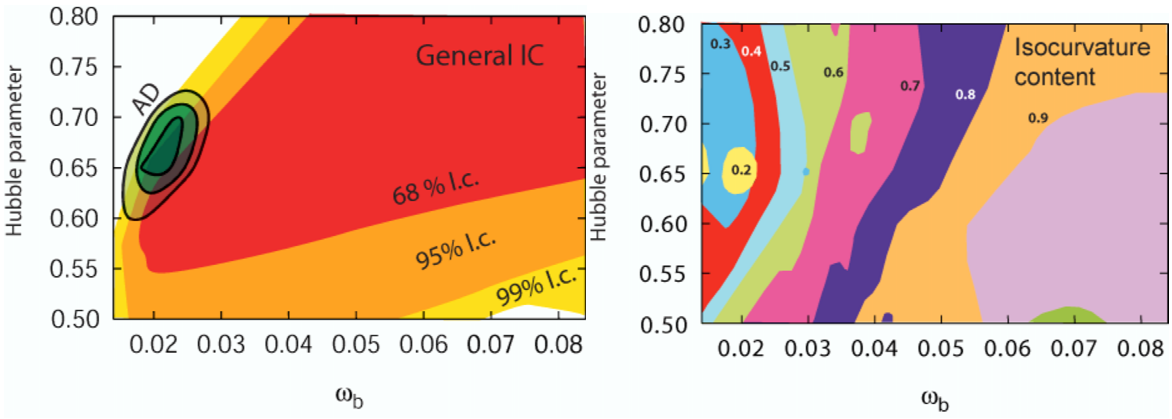

Figure 7.2: Left panel: the contours of 68%, 95%, 99% likelihood content in the $(\omega_{\rm b}, h)$ plane for purely adiabatic models (shadows of green, smaller contours) and for mixed models (red to yellow, large contours). The likelihoods are obtained by maximizing the nuisance parameters, and the initial conditions correlation matrix $\mathbf{M}$ for mixed (i.e. general isocurvature) models. For mixed models, the lowest χ^2 corresponds to even higher values of $\omega_{\rm b}$ and h than those shown in the plot. Right panel: the isocurvature content γ defined in (7.5) of the best fit mixed model as function of the parameters $(\omega_{\rm b}, h)$. A larger value for γ indicates a predominance of the isocurvature modes on the adiabatic one.

7.2.2 How important is the assumption of adiabaticity?

We have shown that in allowing for isocurvature perturbations, one can fit very well pre-WMAP CMB data with cosmological parameters which differ considerably from the ones preferred by adiabatic perturbations alone. More importantly, allowing for generic initial conditions, the ranges of cosmological parameters which can fit the CMB anisotropy data open up to an extent to become nearly meaningless. On the other hand, assuming measurements of cosmological parameters from other methods like direct measurements of the Hubble parameter which yield $h \sim 0.65$ and BBN which implies $\omega_{\rm b} \sim 0.02$, we can use the CMB to limit the isocurvature contribution in the initial conditions (or other unconventional features) and thereby learn something about the very early universe, i.e., the inflationary phase which has generated these initial conditions. For cosmological parameters in the range preferred by other CMB independent measurements ($\Omega_\Lambda \sim 0.7$, $\Omega_{\rm m} \sim 0.3$, $h \sim 0.65$, $\omega_{\rm b} \sim 0.02$) the isocurvature contribution in the initial conditions has to be relatively modest ($\gamma \lesssim 0.3$). We have also checked explicitly that, given these cosmological parameters, a purely isocurvature model, i.e. one with $M_{11} = 0$, cannot fit the data.

Finally, and most importantly, our work shows the danger of calling parameter estimation by CMB anisotropy experiments a "parameter measurement" since the results depend so sensitively (and quite unexpectedly) on the underlying model assumptions. We rather consider CMB anisotropies as an excellent tool to test model assumptions or consistency. In the light of these findings, non-CMB measurements of cosmological parameters acquire even more importance. In short, CMB is the ideal tool to investigate the *primordial parameters* for cosmic structure formation (i.e. the initial conditions), while there are many other possibilities to constrain *cosmological parameters* (Ω_X, h, etc), which we have to use in order to

obtain good limits for possible isocurvature perturbations.

As shown in Bucher et al. (2001) and discussed in § 7.4, CMB temperature anisotropies alone, even if measured with optimal precision limited by cosmic variance, do not allow the degeneracy between cosmological parameters and initial conditions to be removed. Polarization measurements represent an additional non-trivial means to lift this degeneracy and might constrain the contribution of the isocurvature modes to about 10% accuracy (Bucher et al., 2001). The main reason for this is that polarization is mostly sensitive to the quadrupole of the photon distribution rather than the photon density perturbation, these two quantities depending in a different way on the initial conditions. In the same vein, using the normalization of the matter power spectrum (provided it can be measured accurately) also helps to break some of the degeneracies induced by the isocurvature modes, as we show in the next section.

7.3 The cosmological constant problem

Ever since the beginning of modern cosmology, one of the most enigmatic ingredients has been the cosmological constant. Einstein (1917) introduced it to find static cosmological solutions (which are, however, unstable). Later, when the expansion of the Universe had been established, he reportedly called it his "greatest blunder". In relativistic quantum field theory, for symmetry reasons the vacuum energy momentum tensor is of the form $\epsilon g_{\mu\nu}$ for some constant energy density ϵ. The quantity $\Lambda = 8\pi G\epsilon$ can be interpreted as a cosmological constant. Typical values of ϵ expected from particle physics come, for example, from the super-symmetry breaking scale which is expected to be of the order of $\epsilon \gtrsim 1$ TeV^4 leading to $\Lambda \gtrsim 1.7\times10^{-26}$ GeV^2, and corresponding to $\Omega_\Lambda \gtrsim 10^{58}$. Recall that for the density parameter $\Omega_\Lambda \equiv \epsilon/\rho_{\rm crit} = \Lambda/(8\pi G\rho_{\rm crit})$, where $\rho_{\rm crit} = 8.1\times10^{-47}\,h^2$ GeV^4 is the critical density and the fudge factor h is defined by $H_0 = 100\,h$ km s^{-1} Mpc^{-1}, lying in the interval $0.5 \lesssim h \lesssim 0.8$. H_0 is the Hubble parameter today.

Such a result is clearly in contradiction with kinematical observations of the expansion of the universe, which tell us that the value of $\Omega_{\rm tot}$, the density parameter for the total matter-energy content of the universe, is of the order of unity, $\mathcal{O}(\Omega_{\rm tot}) \sim 1$. For a long time, this apparent contradiction has been accepted by most cosmologists and particle physicists, convinced that there must be some deep, not yet understood reason that vacuum energy — which is not felt by gauge-interactions — does not affect the gravitational field either, and hence we measure effectively $\Lambda = 0$. This slightly unsatisfactory situation became really disturbing in 1998, as two groups, which had measured luminosity distances to type Ia supernovae, independently announced that the expansion of the universe is accelerated in the way expected in a universe dominated by a cosmological constant (Riess et al., 1998; Perlmutter et al., 1999). More recent measurements, which extend to higher redshift, seem to strengthen this conclusion (Tonry et al., 2003; Riess et al., 2004), obtaining values of the order $\mathcal{O}(\Omega_{\rm m}) \sim \mathcal{O}(\Omega_\Lambda) \sim 1$ and cannot be explained by any sensible high energy physics model. Tracking scalar fields or quintessence (Ratra & Peebles, 1988; Wetterich, 1988) and other similar ideas (Ferreira & Joyce, 1997) have been introduced in order to mitigate the smallness problem — i.e., the fact that $\epsilon \sim 10^{-46}$ GeV^4. However, none of those is completely successful and really convincing at the moment, see Straumann (2003); Sahni (2004) for reviews.

7.3.1 Does structure formation need a cosmological constant?

After the supernovae Ia results, cosmologists have found many other data-sets which also require a non-vanishing cosmological constant. The most prominent fact is that CMB anisotropies indicate a flat universe, $\Omega_{\rm tot} = \Omega_{\rm m} + \Omega_\Lambda = 1$, while measurements of clustering of matter, e.g., the galaxy power spectrum, require $\Gamma \equiv h\Omega_{\rm m} \simeq 0.2$. But also CMB data alone, with some reasonable prior on the Hubble parameter, point to $\Omega_\Lambda > 0$ at high significance (Spergel et al., 2003).

This cosmological constant problem is probably the greatest enigma in present cosmology. The supernova results are therefore under detailed scrutiny, and there has been a significant amount of work aiming at finding an alternative explanation for the data, see e.g. Meszaros (2002); Blanchard et al. (2003); Alam et al. (2004). Cosmological observations are usually

very sensitive to systematic errors which are often very difficult to discover. Therefore, in cosmology an observational result is usually accepted by the scientific community only if several independent data-sets lead to the same conclusion. But this seems to be exactly the case for the cosmological constant.

It is therefore imperative to investigate in detail whether present structure formation data does require a cosmological constant, by asking whether enlarging the space of models for structure formation does mitigate the cosmological constant problem. There are several ways to enlarge the model space, e.g. one may allow for features in the primordial power spectrum, like a kink (Barriga et al., 2001). Here we study the cosmological constant problem in relation to the initial conditions for the cosmological perturbations.

In a first step we discuss once more the usual results obtained assuming purely adiabatic models and we investigate the extent to which pre-WMAP CMB data alone or combined with large-scale structure measurements require $\Omega_\Lambda \neq 0$ in a flat universe, presenting the findings published in Trotta et al. (2003). We shall first proceed with the usual Bayesian analysis, but we also discuss the results which are obtained in a frequentist approach. We find that even if $\Omega_\Lambda = 0$ is outside the high likelihood region in a Bayesian approach this is no longer the case from the frequentist point of view. In other words the probability that a model with vanishing Ω_Λ leads to the present-day observed CMB and large-scale structure data is not exceedingly small.

We then study how the results are modified if we allow for general isocurvature contributions to the initial conditions. In this first study of the matter power spectrum from general isocurvature modes we discover that a COBE-normalized matter power spectrum reproduces the observed amplitude only if it is highly dominated by the adiabatic component. Hence the isocurvature modes cannot contribute significantly to the matter power spectrum and do not lead to a degeneracy in the initial conditions for the matter power spectrum when combined with CMB data.

7.3.2 CMB and large scale structure data analysis

The pre-WMAP CMB measurements, from BOOMERanG (Netterfield et al., 2002), MAXIMA (Lee et al., 2001), DASI (Halverson et al., 2002), VSA (Scott et al., 2003; Taylor et al., 2003), CBI (Pearson et al., 2003) and Archeops (Benoit et al., 2003a) are in very good agreement up to the third peak in the angular temperature power spectrum of CMB anisotropies, $\ell \sim 1000$. In our analysis we therefore use the COBE data (Smoot et al., 1992; Bennett et al., 1994) in the decorrelated compilation of Tegmark & Hamilton (1997) (7 points excluding the quadrupole) for the ℓ region $3 \leq \ell \leq 20$ and the BOOMERanG data to cover the higher ℓ part of the spectrum (19 points in the range $100 \leq \ell \leq 1000$). Since Archeops has the smallest error bars in the region of the first acoustic peak, we also include this data-set (16 points in the range $15 \leq \ell \leq 350$). Including any of the other mentioned data does not influence our results significantly. The BOOMERanG and Archeops absolute calibration errors (10% and 7% at 1σ, respectively) as well as the uncertainty of the BOOMERanG beam size are included as additional Gaussian nuisance parameters, and are maximized over. We make use of the Archeops window functions available from the Archeops Website (2003), while for BOOMERanG a top-hat window is assumed. For the matter power spectrum, we use the galaxy-galaxy power spectrum from the 2dF data which is obtained from the redshift of

about 10^5 galaxies (Tegmark et al., 2002). We include only the 22 decorrelated points in the linear regime, i.e., in the range $0.017 \leq k \leq 0.314 \quad [h\ \mathrm{Mpc}^{-1}]$, and the window functions of Tegmark et al. (2002) which can be found at Tegmark's Website (2003).

Our grid of models is restricted to flat universes and we assume purely scalar perturbations. Since the goal here is more to make a conceptual point than to consider the most generic model, we fix the baryon density to the BBN preferred value $\Omega_{\rm b}h^2 \equiv \omega_{\rm b} = 0.020$ (Burles et al., 2001) and we investigate the following 3-dimensional grid in the space of cosmological parameters:

$$\begin{aligned} 0.35 < h &< 1.00 \quad (0.025)\,, \\ 0.00 < \Omega_\Lambda &< 0.95 \quad (0.05)\,, \\ 0.80 < n_{\rm s} &< 1.20 \quad (0.05)\,, \end{aligned} \tag{7.6}$$

where $n_{\rm s}$ is the scalar spectral index, which again we take to be the same for all modes, and the numbers in parenthesis give the step size we use. The total matter content $\Omega_{\rm m} \equiv \Omega_{\rm cdm} + \Omega_{\rm b}$ is $\Omega_{\rm m} = 1 - \Omega_\Lambda$, and $\Omega_{\rm cdm}$ indicates the cold dark matter contribution. For all models the optical depth of reionization is $\tau = 0$ and we have three families of massless neutrinos. For each grid point we compute the ten CMB and matter power spectra, one for each independent set of initial conditions, as explained in § 4.3. The initial condition correlation matrix $\mathbf{M}$ is parameterized using the ten dimensional hypercube parameters presented on page 108.

For a given initial conditions correlation matrix $\mathbf{M}$ and spectral index $n_{\rm s}$, we quantify the isocurvature contribution to the CMB temperature anisotropy by the phenomenological parameter β defined as

$$\beta \equiv \frac{\sum_{X=\mathrm{CI,NV,ND}} \left\langle (\ell(\ell+1)) C_\ell^{(X,X)} \right\rangle_\ell}{\sum_{Y=\mathrm{AD,CI,NV,ND}} \left\langle \ell(\ell+1) C_\ell^{(Y,Y)} \right\rangle_\ell}, \tag{7.7}$$

where the average $\langle\cdot\rangle$ is taken in the ℓ range of interest, in our case $3 \leq \ell \leq 1000$, and where $C_\ell^{(X,X)}$ stands for the auto-correlator of the CMB anisotropies with initial conditions X. This quantity measures the average power of the adiabatic and isocurvature modes over the full multipole range, and therefore it gives a more phenomenological description of the isocurvature contribution than the parameter γ used in the previous section, and defined in Eq. (7.5, page 188).

As highlighted in § 5.1.5, the correct interpretation of Bayesian statistics is in terms of most likely regions in parameter space, while the frequentist approach is required in order to obtain exclusion intervals for the parameters. In order to answer the question of whether the CMB and large scale structure data exclude with a given confidence the value $\Omega_\Lambda = 0$, we use the frequentist statistics, and compare the result with the usual Bayesian approach.

7.3.3 Adiabatic perturbations

We first fit CMB data only ($N = 42$) by maximizing $M = 7$ parameters, i.e., the three nuisance parameters, $n_{\rm s}$, h, Ω_Λ and the overall amplitude of the adiabatic spectrum, and we find (Bayesian likelihood intervals on Ω_Λ alone):

$$\Omega_\Lambda = 0.80^{+0.10}_{-0.35} \text{ at } 2\sigma \qquad \text{and} \qquad {}^{+0.12}_{-0.80} \text{ at } 3\sigma. \tag{7.8}$$

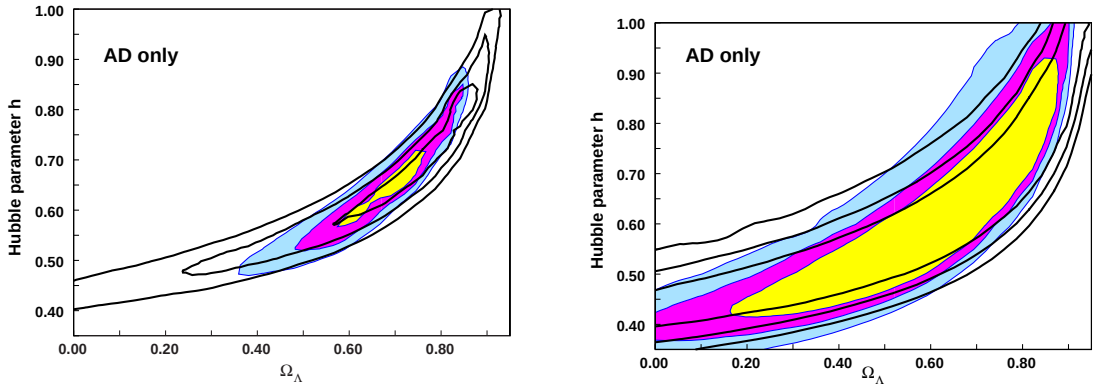

Figure 7.3: Joint likelihood contours (Bayesian, left panel) and confidence contours (frequentist, right panel), with CMB only (solid lines,1σ, 2σ, 3σ contours) and CMB+2dF (filled) for purely adiabatic initial conditions. In the right panel, the number of effective degrees of freedom is $F_{\rm eff} = 31$ for CMB alone $F_{\rm eff} = 50$ for CMB+2dF.

The asymmetry in the intervals arises because the value of Ω_Λ for our maximum likelihood (ML) model is relatively large. One could achieve a better precision in determining the ML value of Ω_Λ by using a finer grid and varying $\omega_{\rm b}$ as well, which has extensively been done in the literature and is not the scope of this work. Moreover, the position of the acoustic peaks in CMB anisotropies is mostly sensitive to the age of the universe at recombination, which depends only on $\Omega_{\rm m}h^2$, and to the angular diameter distance, which depends on $\Omega_{\rm m}$, Ω_Λ and the curvature of the universe. When the universe is flat, the angular diameter distance is weakly dependent on the relative amounts of $\Omega_{\rm m}$ and Ω_Λ as long as Ω_Λ is not too large, see § 4.1.2 and Fig. 4.1 on page 96. Hence, one can achieve a sufficiently low value of $\Omega_{\rm m}h^2$ either via a large cosmological constant or via a very low Hubble parameter, $h \lesssim 0.45$.

We now include the matter power spectrum $P_{\rm m}$, assuming $P_{\rm m} = b^2 P_{\rm g}$, where $P_{\rm g}$ is the observed galaxy power spectrum and b some unknown bias factor (assumed to be scale independent), over which we maximize. Inclusion of this data in the analysis breaks the Ω_Λ, h degeneracy, since $P_{\rm m}$ is mainly sensitive to the shape parameter $\Gamma \equiv \Omega_{\rm m}h$. We therefore obtain significantly tighter overall likelihood intervals for Ω_Λ:

$$\Omega_\Lambda = 0.70^{+0.13}_{-0.17} \text{ at } 2\sigma \qquad \text{and} \qquad {}^{+0.15}_{-0.27} \text{ at } 3\sigma\,. \tag{7.9}$$

We plot joint likelihood contours (Bayesian) for Ω_Λ, h with purely adiabatic initial conditions in the left panel of Fig. 7.3. From the Bayesian analysis, one concludes that CMB and 2dF together require a non-zero cosmological constant at very high significance, more than 7σ for the points in our grid! Note that the ML point has a reduced chi-square $\hat{\chi}^2_{F=56} = 0.59$, significantly less than unity.

The frequentist analysis, however, excludes a much smaller region of parameter space, cf. the right panel of Fig. 7.3. The frequentist contours must be drawn for the effective number of degrees of freedom, i.e., using the number of effectively independent data points. We can therefore roughly take into account a 10% correlation, which is the maximum correlation between data points given in Netterfield et al. (2002); Benoit et al. (2003a), by replacing F

by the effective number of degrees of freedom, $F_{\rm eff} = 0.9N - M$, and rounding to the next larger integer (to be conservative). One could argue that the BOOMERanG and Archeops data points are not completely independent, since BOOMERanG observed a portion of the same sky patch as measured by Archeops. This possible correlation is difficult to quantify, but should not be too important since the sky portion observed by Archeops is a factor of 10 larger than BOOMERanG's and therefore we ignore it here. The right panel of Fig. 7.3 is drawn with $F_{\rm eff} = 31$ for CMB alone and $F_{\rm eff} = 50$ for CMB+2dF, but we have checked that our results do not change much if we use a 5% correlation.

It is interesting to note that there are regions in the left panel which are excluded with a certain confidence by CMB data alone but are no longer excluded at the same confidence when we include the 2dF data. In other words, it would seem that taking into account more data and therefore more knowledge about the universe, does not systematically exclude more models, i.e., the CMB+2dF contours are not always contained in the CMB alone contours. This apparent contradiction vanishes when one realizes that the confidence limits on, e.g., Ω_Λ alone in the frequentist approach are just the projection of the confidence contours of the right panel on the Ω_Λ axis. One can readily verify in the right panel that the confidence limits for the combined data-set are always smaller than the ones for CMB data alone. There are points with $\Omega_\Lambda = 0$ and $h \simeq 0.40$ which are still compatible within 2σ with both 2dF and CMB data, at the price of pushing somewhat the other parameters. In the best fit with $\Omega_\Lambda = 0$ shown in Fig. 7.4, one has to live with a red spectral index $n_{\rm s} = 0.80$. Furthermore, the calibration of the BOOMERanG and Archeops data points is reduced in this fit by 34% and 26%, respectively, i.e., more than 3 times the quoted 1σ systematic error.

In both cases, it is clear that one can exploit the Ω_Λ, h degeneracy to fit CMB data alone with a model having $\Omega_\Lambda = 0$. For a flat universe like the one we are considering, one has then to use a much smaller value of the Hubble parameter than the one indicated by other measurements, most notably the HST Key Project (Freedman et al., 2001), which gives $h = 0.72 \pm 0.08$. The 2dF data are mainly sensitive to the shape parameter $\Gamma \sim 0.2$, hence 2dF with $\Omega_{\rm m} = 1.0$ would require an even lower value of h which is not compatible with CMB. Therefore inclusion of 2dF data tends to exclude any flat model without a cosmological constant. Summing up, for purely adiabatic initial conditions the Bayesian approach gives very strong support to $\Omega_\Lambda \neq 0$; in the more conservative frequentist point of view, while $\Omega_\Lambda \neq 0$ cannot be excluded with very high confidence, the combination of 2dF and pre-WMAP CMB data start to be incompatible with a flat universe with vanishing cosmological constant. These conclusions are in qualitative agreement with previous works using comparable data (Netterfield et al., 2002; Pryke et al., 2002; Lewis & Bridle, 2002; Wang et al., 2002; Durrer et al., 2003b; Rubino-Martin et al., 2003; Benoit et al., 2003b). In the next section we investigate the stability of these well known results with respect to inclusion of non-adiabatic initial conditions.

7.3.4 Mixed adiabatic and isocurvature perturbations

We now enlarge the space of models by including all possible isocurvature modes with arbitrary correlations among themselves and the adiabatic mode as described in the previous section, but with the restriction that all modes have the same spectral index. We first consider CMB data only and maximize over initial conditions. The number of parameters increases

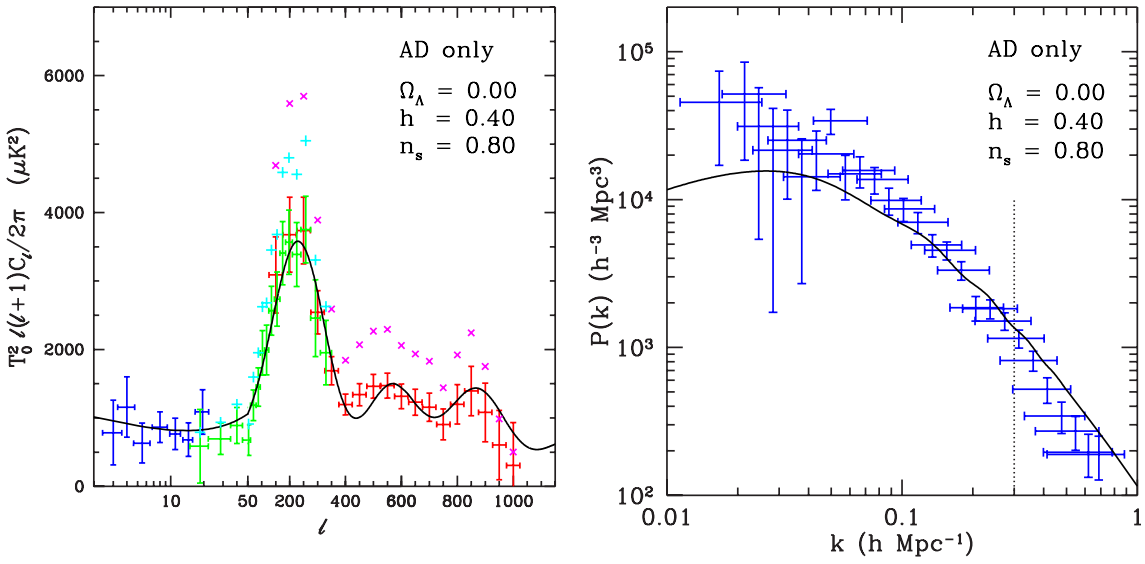

Figure 7.4: Best fit with $\Omega_\Lambda = 0$ and purely adiabatic initial conditions, compatible with CMB and 2dF data within 2σ confidence level (frequentist). In the right panel, only the 2dF data points left of the vertical, dotted line — i.e., in the linear region — have been included in the analysis. Note the low CMB first acoustic peak in the left panel due to the joint effect of the red spectral index and of the absence of early ISW effect. In this fit, the calibration of BOOMERanG (red/dark gray errorbars) and Archeops (green/light gray errorbars) has been reduced by 34% and 26%, respectively. To appreciate the difference, we plot the non recalibrated value of the BOOMERanG and Archeops data points as diagonal/magenta crosses and vertical/light blue crosses, respectively. Even though the fit is "by eye" very good, it seems highly unlikely that the calibration error is so large.

by nine and the number of degrees of freedom decreases correspondingly with respect to the purely adiabatic case considered above.

Likelihood (Bayesian, left panel of Fig. 7.5) and confidence (frequentist, right panel of Fig. 7.5) contours widen up somewhat along the degeneracy line. The enlargement is less dramatic than in the case of the baryon density presented in § 7.2. This is partially due to our prior of flatness which reduces the space of models to those which are almost degenerate in the angular diameter distance. Most of our models have the first acoustic peak of the adiabatic mode already in the region preferred by experiments, hence in most of the fits, isocurvature modes play a modest role, especially in the parameter regions with large Ω_Λ, h (cf. Fig. 7.9 and the discussion below). Nevertheless, because of the Ω_Λ, h degeneracy, even a modest widening of the contours along the degeneracy line results in an important enlargement of the likelihood limits. The ML point does not depart very much from the purely adiabatic case, but now we cannot constrain Ω_Λ at more than 1σ (Bayesian, CMB only):

$$\Omega_\Lambda = 0.85^{+0.05}_{-0.35} \text{ at } 1\sigma \quad , \tag{7.10}$$

and no limits for $0.0 \leq \Omega_\Lambda \leq 0.95$ at higher confidence.

In Fig. 7.6 we plot the dark matter power spectra of the different auto- (left panel)

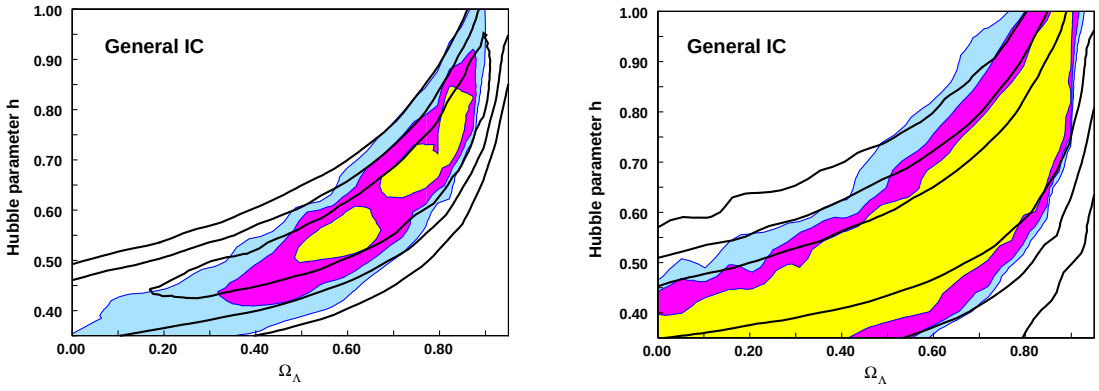

Figure 7.5: Joint likelihood contours (Bayesian, left panel) and confidence contours (frequentist, right panel), with CMB only (solid lines) and CMB+2dF (filled) after maximization over general isocurvature initial conditions. The likelihood/probability content is 1σ, 2σ, 3σ, from the center to the outside. The disconnected 1σ region in the left panel is an artificial feature due to the grid resolution. In the right panel, the number of effective degrees of freedom is $F_{\rm eff} = 22$ for CMB alone $F_{\rm eff} = 41$ for CMB+2dF.

and cross-correlators (right panel) for a concordance model. The norm of each pure mode (AD, CI, ND, NV) is chosen such that the corresponding CMB power spectrum is COBE-normalized. The cross-correlators are normalized according to totally correlated spectra, i.e.

$$M_{(\mathrm{X,Y})} = \sqrt{M_\mathrm{X} M_\mathrm{Y}/2}\,, \tag{7.11}$$

where $M_{(\mathrm{X,Y})}$ denotes the norm of the cross-correlator between the modes X,Y and M_X the norm of the pure mode X. A crucial result is that the COBE-normalized amplitude of the adiabatic matter power spectrum is nearly two orders of magnitude larger than the isocurvature contribution. The main reason for this is the amplitude of the Sachs-Wolfe plateau which is about $\frac{1}{3}\Phi$ for adiabatic perturbations and 2Φ for isocurvature perturbations, where Φ is the gravitational potential at last scattering, see Eq. (4.3) and Eq. (4.4, page 91). This difference of a factor of about 36 in the power spectrum on large scales is clearly visible in the comparison of P_AD and P_CI (the difference increases at smaller scales). The case of the neutrino modes is even worse since they start with vanishing dark matter perturbations. That the CDM isocurvature matter power spectrum is much lower than the adiabatic one has been known for some time (see e.g. Stompor et al., 1996; Pierpaoli et al., 1999). However, it was not recognized before that the same holds true for the neutrino isocurvature matter power spectra as well, and – more importantly – that this leads to a way to break the strong degeneracy among initial conditions which is present in the CMB power spectrum alone.

In an analysis with general initial conditions including the 2dF data only we obtain very broad likelihood and confidence contours which exclude only the lower right corner of the (Ω_Λ, h) plane. In contrast to the CMB power spectrum, the matter power spectrum can be fitted with extremely high adiabatic and isocurvature contributions, which are then typically cancelled by large anti-correlations between the spectra. This behavior is exemplified for a model with general isocurvature initial conditions and $\Omega_\Lambda = 0.70$, $h = 0.65$, $n_\mathrm{s} = 1.0$

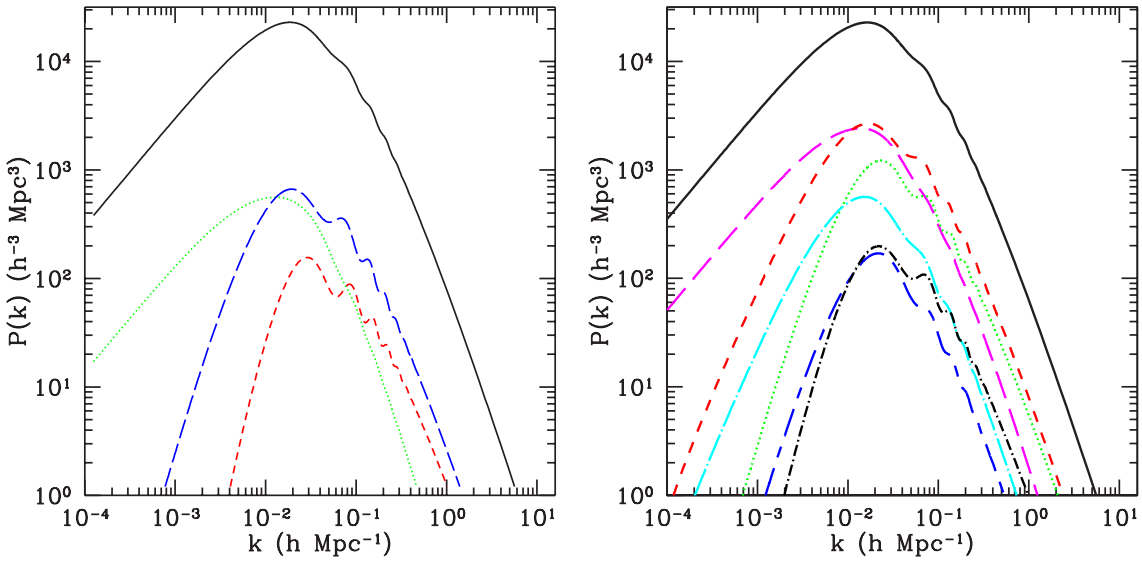

Figure 7.6: Dark matter power spectra of the different auto- (left panel) and cross-correlators (right panel) for a concordance model with $\Omega_\Lambda = 0.70$, $h = 0.65$, $n_s = 1.0$, $\omega_b = 0.020$, with the corresponding CMB power spectrum COBE-normalized. The color and line style codes are as follows: in the left panel, adiabatic (AD): solid/black line; CDM isocurvature (CI): dotted/green line; neutrino density (ND): short-dashed/red line; neutrino velocity (NV): long-dashed/blue line; in the right panel, AD: solid/black line (for comparison), $\langle \mathrm{AD}, \mathrm{CI} \rangle$: long-dashed/magenta line, $\langle \mathrm{AD}, \mathrm{ND} \rangle$: dotted/green line, $\langle \mathrm{AD}, \mathrm{NV} \rangle$: short-dashed/red line, $\langle \mathrm{CI}, \mathrm{ND} \rangle$: dot-short dashed/blue line, $\langle \mathrm{CI}, \mathrm{NV} \rangle$: dot-long dashed/light-blue line, and $\langle \mathrm{ND}, \mathrm{NV} \rangle$: dot-short dashed/black line. The adiabatic mode is by far dominant over all others.

in Fig. 7.7. The best fits with 2dF data only are dominated by large isocurvature cross-correlations. Clearly, the resulting CMB power spectrum is highly inconsistent with the COBE data. Hence such "bizarre" possibilities are immediately ruled out once we include CMB data. Conversely, moderate isocurvature contributions can help fitting the CMB data, and do not influence the matter power spectrum, which is completely dominated by the adiabatic mode alone.

Combining CMB and 2dF data we find now (Bayesian, mixed isocurvature models):

$$\Omega_\Lambda = 0.65^{+0.22}_{-0.25} \text{ at } 2\sigma \qquad \text{and} \qquad {}^{+0.25}_{-0.48} \text{ at } 3\sigma\,. \tag{7.12}$$

The likelihood limits are larger than for the purely adiabatic case but it is interesting that the Bayesian analysis still excludes $\Omega_\Lambda = 0$ at more than 3σ even with general initial conditions, for the class of models considered here. Because of the above explained reason, the widening of the limits is not as drastic as one might fear. Therefore, combination of CMB and LSS measurements turn out to be an ideal tool to constrain the isocurvature contribution to the initial conditions.

From the frequentist point of view, one notices that the region in the Ω_Λ, h plane which is incompatible with data at more than 3σ is nearly independent on the choice of initial conditions (compare the right panels of Fig. 7.3 and Fig. 7.5). Enlarging the space of initial

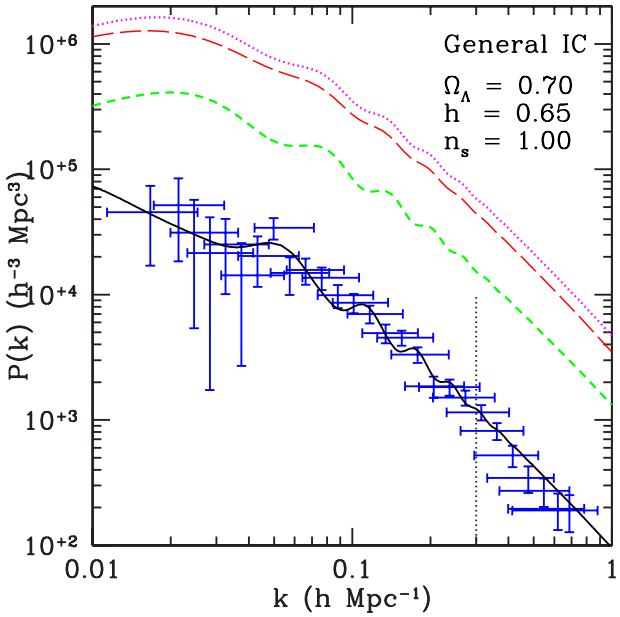

Figure 7.7: Concordance model fit with general isocurvature initial conditions and 2dF data only. The total spectrum (solid/black) is the result of a large cancellation of the purely adiabatic part (long-dashed/red) by the large, negative sum of the various correlators (dotted/magenta, plotted in absolute value). The short-dashed/green curve is the sum of the three pure isocurvature modes, CI, ND and NV. Note that the resulting total spectrum is less than one tenth of the purely adiabatic part.

conditions seemingly does not have a relevant benefit on fitting CMB and 2dF data with or without a cosmological constant. The reason for this is that the (COBE-normalized) matter power spectrum is dominated by its adiabatic component and therefore the requirement $\Omega_{\rm m}h \sim 0.2$ remains valid. In Fig. 7.8 we plot the best fit model with general initial conditions and $\Omega_\Lambda = 0$. We summarize our likelihood and confidence intervals on Ω_Λ (this parameter only) in Table 7.1.

In Fig. 7.9 we plot the isocurvature contribution to the best fit models with CMB and 2dF in terms of the parameter β defined in (7.7). The best fit with $\Omega_\Lambda = 0$ has an isocurvature contribution of about 40%. We can put a constraint on the maximal isocurvature contribution allowed by combining this plot with the exclusion plot obtained with the frequentist approach, Fig. 7.5 right panel. The result is that frequentist statistics limits the isocurvature content β to be

$$\beta \lesssim 0.4 \quad (2\sigma \text{ c.l.}). \tag{7.13}$$

7.3.5 Do isocurvature perturbations mitigate the Λ problem?

There are three main conclusions we can draw from these results. The first one is not new, but seems to be dangerously forgotten in recent cosmological parameters estimation literature: namely that likelihood contours cannot be used as "exclusion plots". The latter are usually

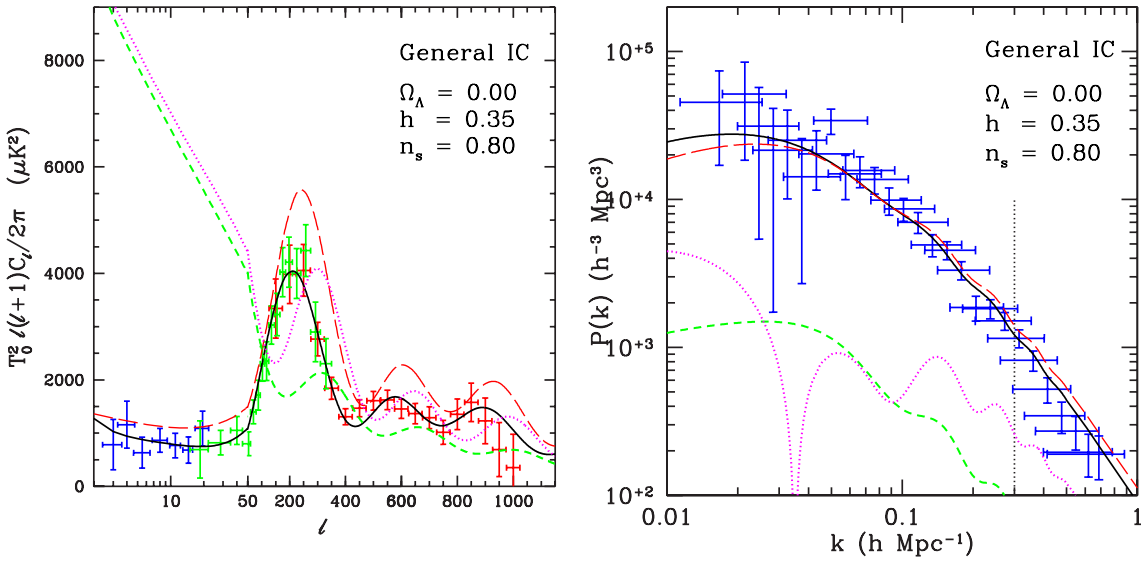

Figure 7.8: Best fit with general isocurvature models and $\Omega_\Lambda = 0$. As for the purely adiabatic case, even with general initial conditions the absence of the cosmological constant suppresses in an important way the height of the first peak. In both panels we plot the best total spectrum (solid/black), the purely adiabatic contribution (long-dashed/red), the sum of the pure isocurvature modes (short-dashed/green) and the sum of the correlators (dotted/magenta, multiplied by -1 in the left panel and in absolute value in the right panel). The matter power spectrum is completely dominated by the adiabatic mode, while the correlators play an important role in cancelling unwanted contributions in the CMB power spectrum at the level of the first peak and especially in the COBE region. For this model we have an isocurvature content $\beta = 0.39$, while the BOOMERanG and Archeops calibrations are reduced by 28% and 12%, respectively. The color codes for the error-bars are the same as in Fig. 7.4.

substantially wider, less stringent. A more rigorous possibility are frequentist probabilities, which however suffer from the dependence on the number of really independent measurements which is often very difficult to come by.

Secondly, we have found that in COBE-normalized fluctuations, the matter power spectrum has negligible isocurvature contributions and is essentially given by the adiabatic mode. Hence the shape of the observed matter power spectrum still requires $\Omega_{\rm m}h \simeq 0.2$, independent of the choice of initial conditions. Due to this behavior, the condition $\Omega = \Omega_\Lambda + \Omega_{\rm m} = 1$ requires either a cosmological constant or a very small value for the Hubble parameter, independently from the isocurvature contribution to the initial conditions.

The third conclusion concerns the presence of a cosmological constant from pre-WMAP CMB data combined with the 2dF matter power spectrum: For flat models, a likelihood (Bayesian) analysis strongly favors a non-vanishing cosmological constant. Even if we allow for isocurvature contributions with arbitrary correlations, a vanishing cosmological constant is still outside the 3σ likelihood range. It is possible that there are open models, which we did not consider here, in which the NV mode would be dominant,: this because it presents a first acoustic peak at $\ell = 170$ in flat models, which would be displaced to a larger multipole value, as preferred by data, in an open Universe, thereby possibly giving a good fit to CMB data

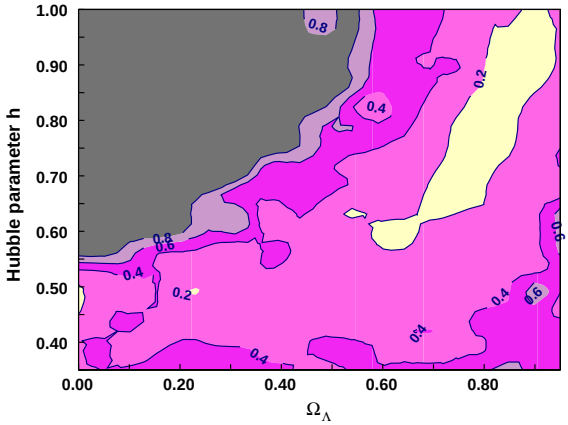

Figure 7.9: Isocurvature content $0.0 \leq \beta \leq 1.0$ of best fit models with CMB and 2dF data. The contours are for $\beta = 0.20, 0.40, 0.60, 0.80$ from the center to the outside.

Purely adiabatic									
		Bayesian[1]			Frequentist[2]				
Data-sets	Ω_Λ	1σ	2σ	3σ	1σ	2σ	3σ	F	χ^2/F
CMB	0.80	$^{+0.08}_{-0.08}$	$^{+0.10}_{-0.35}$	$^{+0.12}_{-}$	< 0.93	–	–	35	0.58
CMB +2dF	0.70	$^{+0.05}_{-0.05}$	$^{+0.13}_{-0.17}$	$^{+0.15}_{-0.27}$	$\Omega_\Lambda{}^{<0.90}_{>0.15}$	< 0.92	< 0.92	56	0.59
General isocurvature									
CMB	0.85	$^{+0.05}_{-0.35}$	–	–	–	–	–	26	0.74
CMB+2dF	0.65	$^{+0.15}_{-0.10}$	$^{+0.22}_{-0.25}$	$^{+0.25}_{-0.48}$	< 0.90	< 0.92	< 0.95	47	0.67

[1] Likelihood interval.

[2] Region not excluded by data with given confidence.

Table 7.1: Likelihood (Bayesian) and confidence (frequentist) intervals for Ω_Λ alone (all other parameters maximized). A bar, $-$, indicates that at the given likelihood/confidence level the analysis cannot constraint Ω_Λ in the range $0.0 \leq \Omega_\Lambda \leq 0.95$. Where the quoted interval is smaller than our grid resolution, an interpolation between models has been used.

and allow for the observed shape parameter Γ with a reasonable value of h. This question remains to be investigated in detail.

The situation changes considerably in the frequentist approach. There, even for purely adiabatic models, $\Omega_\Lambda = 0$ is still within 3σ for a value of $h \leq 0.48$ which is marginally defendable. The conclusion does not change very much when we allow for generic initial conditions.

7.4 Precision cosmology independent of initial conditions

As we have seen, it is difficult to simultaneously constrain both the type of initial conditions and the cosmological parameters using CMB alone. The future high accuracy measurements of CMB polarization will help substantially in breaking degeneracies between initial conditions. The degeneracies in the parameter dependence of temperature and polarization are almost orthogonal, and polarization can therefore lift "flat directions" in parameter space.

To determine cosmological parameters independently on the initial conditions, one includes general isocurvature modes, and then marginalize over them. Bucher et al. (2002, 2001) considered forecasts for WMAP and Planck, and found that admitting isocurvature modes would ruin the ability of WMAP to determine the cosmological parameters with temperature information only. They also highlighted that polarization measurements would be decisive in assisting into the reconstruction of the cosmological parameters when allowing for general isocurvature initial conditions. Their results were obtained with a Fisher matrix analysis on a cosmological parameter set which, according to Kosowsky et al. (2002), leads to large overestimates of the expected errors. We have reproduced their study (Trotta & Durrer, 2004), using for the Fisher matrix forecast the normal parameter set described in § 4.2 so that we obtain forecasts not for the highly degenerate directions defined by the cosmological parameters, but rather for orthogonal combinations which are well measured by the CMB. Along these directions, forecasts are much more reliable. The main features are summarized in Fig. 7.10, where we plot the expected 1σ error in percent for the six quantities which are directly probed by the CMB with good accuracy (see figure caption). We omit the energy density in the cosmological constant, which is ill-determined with CMB alone because of the

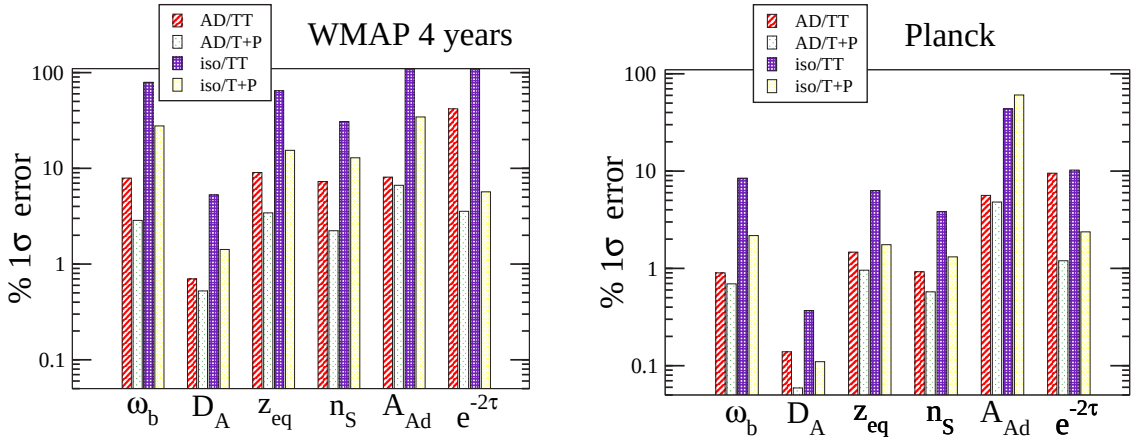

Figure 7.10: Fisher matrix forecast for the percent 1σ errors on six quantities which are well determined by CMB alone with and without inclusion of general isocurvature initial conditions. The left (right) panel is a forecast for WMAP four year mission (Planck). From left to right, on the abscissa axis: the baryon density, ω_b, the angular diameter distance D_A, the redshift of matter-radiation equality $z_{\rm eq}$, the scalar spectral index $n_{\rm s}$, the scalar adiabatic amplitude $A_{\rm Ad}$ and a function of the optical depth to reionization, $\tau_{\rm re}$. In the legend, "AD" means that only adiabatic fluctuations were included, "iso" means that general isocurvature modes were included and marginalized over. "TT" uses temperature information alone, "T+P" has temperature, E-T correlation and E-polarization.

geometrical degeneracy. We do not restrict our analysis to flat models, but include spaces with non-zero curvature.

For WMAP the errors on normal parameters will increase roughly by a factor ten with respect to the purely adiabatic scenario if one marginalizes over general initial conditions, when temperature information alone is considered (cf. first and third bar in the left panel). When the full polarization information is included, however, the errors will still be within approximately 10 to 30% even in the general isocurvature scenario. From the right panel, we deduce that for the Planck experiment the worsening of the errors will be much less if the high quality polarization information is included. Roughly speaking, by including isocurvature modes we expect errors which are larger than in the adiabatic case by about a factor of two, but mostly still within the few percent accuracy. These findings are in qualitative agreement with Bucher et al. (2001), while providing a quantitatively more reliable estimate of the expected accuracy.

This shows that the CMB alone will be able to provide high precision cosmology even if the strong assumption of purely adiabatic initial conditions will be relaxed. Combining CMB results with other observation which independently constrain the cosmological parameters, will enable us to fully open this window to the mysterious epoch of the very early universe.

Publication list

---

1. **Trotta, R.** & Durrer, R. (2004). Testing the paradigm of adiabaticity. In Ruffini et al., editors, *Proceedings of the X Marcel Grossman Meeting, 20-26 July 2003, Rio de Janeiro*, astro-ph/0402032. To appear.

2. Rocha, G.,**Trotta, R.**, Martins, C., Melchiorri, A., Avelino, P., Bean, R., & Viana, P. (2004). Measuring α in the early Universe: CMB polarization, reionization and the Fisher matrix analysis. *Mon. Not. Roy. Astron. Soc.*, 352:20–38, astro-ph/0309211.

3. **Trotta, R.** & Hansen, S. H. (2004). Observing the helium abundance with CMB. *Phys. Rev.*, D69:023509, astro-ph/0306588.

4. Rocha, G.,**Trotta, R.**, Martins, C., Melchiorri, A., Avelino, P., Bean, R., & Viana, P. (2003). New constraints on varying α *New Astronomy Reviews*, Proceedings of the 2nd CMBNET Meeting, Oxford, 20-21 February 2003, Oxford, UK, 47: 863-869, astro-ph/0309205.

5. **Trotta, R.** (2003). The cosmological constant and the paradigm of adiabaticity. *New Astronomy Reviews*, Proceedings of the 2nd CMBNET Meeting, Oxford, 20-21 February 2003, Oxford, UK, 47: 769-774, astro-ph/0304525.

6. Martins, C. J. A. P., Melchiorri, A., Rocha, G.,**Trotta, R.**, Avelino, P., & Viana, P. (2004). WMAP constraints on varying α and the promise of reionization. *Phys. Lett. B*, 585:29, astro-ph/0302295.

7. **Trotta, R.**, Riazuelo, A., & Durrer, R. (2003). The cosmological constant and general isocurvature initial conditions. *Phys. Rev.*, D67:063520, astro-ph/0211600.

8. Martins, C. J. A. P., Melchiorri, A.,**Trotta, R.**, Bean, R., Rocha, G., Avelino, P., & Viana, P. (2002). Measuring α in the early universe: CMB temperature, large-scale structure and Fisher matrix analysis. *Phys. Rev.*, D66:023505, astro-ph/0203149.

9. Bowen, R., Hansen, S. H., Melchiorri, A., Silk, J., & **Trotta, R.** (2002). The impact of an extra background of relativistic particles on the cosmological parameters derived from microwave background anisotropies. *Mon. Not. Roy. Astron. Soc.*, 334:760, astro-ph/0110636.

10. **Trotta, R.**, Riazuelo, A., & Durrer, R. (2001). Reproducing cosmic microwave background anisotropies with mixed isocurvature perturbations. *Phys. Rev. Lett.*, 87:231301, astro-ph/0104017.

Bibliography

Abazajian, K. N., Beacom, J. F., & Bell, N. F. (2002). Stringent constraints on cosmological neutrino antineutrino asymmetries from synchronized flavor transformation. *Phys. Rev.*, D66:013008, astro-ph/0203442.

Abramowitz, M. & Stegun, I. A. (1970). *Handbook of Mathematical functions*. Dover Publications, 9th edition.

Abroe, M. E. et al. (2002). Frequentist Estimation of Cosmological Parameters from the, MAXIMA-1 Cosmic Microwave Background Anisotropy Data. *Mon. Not. Roy. Astron. Soc.*, 334:11, astro-ph/0111010.

Alam, U., Sahni, V., & Starobinsky, A. A. (2004). The case for dynamical dark energy revisited. *JCAP*, 0406:008, astro-ph/0403687.

Albrecht, A. & Skordis, C. (2000). Phenomenology of a realistic accelerating universe using only Planck-scale physics. *Phys. Rev. Lett.*, 84:2076–2079, astro-ph/9908085.

Amendola, L., Gordon, C., Wands, D., & Sasaki, M. (2002). Correlated perturbations from inflation and the cosmic microwave background. *Phys. Rev. Lett.*, 88:211302, astro-ph/0107089.

Archeops Website (2003). `http://www.archeops.org/`.

Avelino, P. P. et al. (2001). Early-universe constraints on a time-varying fine structure constant. *Phys. Rev.*, D64:103505, astro-ph/0102144.

Avelino, P. P., Martins, C. J. A. P., Rocha, G., & Viana, P. (2000). Looking for a varying α in the Cosmic Microwave Background. *Phys. Rev.*, D62:123508, astro-ph/0008446.

Bahcall, N. A., Cen, R., Dave, R., Ostriker, J. P., & Yu, Q. (2000). The Mass-To-Light Function: Antibias and Ω_m. *Astrophys. J.*, 541, astro-ph/0002310.

Bahcall, N. A. et al. (2003). The cluster mass function from early SDSS data: Cosmological implications. *Astrophys. J.*, 585:182–190, astro-ph/0205490.

Bardeen, J. M. (1980). Gauge invariant cosmological perturbations. *Phys. Rev.*, D22:1882–1905.

Barger, V., Kneller, J. P., Langacker, P., Marfatia, D., & Steigman, G. (2003a). Hiding relativistic degrees of freedom in the early universe. *Phys. Lett.*, B569:123–128, hep-ph/0306061.

Barger, V., Kneller, J. P., Lee, H.-S., Marfatia, D., & Steigman, G. (2003b). Effective number of neutrinos and baryon asymmetry from BBN and WMAP. *Phys. Lett.*, B566:8–18, hep-ph/0305075.

Barger, V. D. & Kao, C. (2001). Implications of new CMB data for neutralino dark matter. *Phys. Lett.*, B518:117–122, hep-ph/0106189.

Barriga, J., Gaztanaga, E., Santos, M. G., & Sarkar, S. (2001). On the APM power spectrum and the CMB anisotropy: Evidence for a phase transition during inflation? *Mon. Not. Roy. Astron. Soc.*, 324:977, astro-ph/0011398.

Barrow, J. D., Sandvik, H. B., & Magueijo, J. (2002). The behaviour of varying-alpha cosmologies. *Phys. Rev.*, D65:063504, astro-ph/0109414.

Bartelmann, M. & Schneider, P. (2001). Weak gravitational lensing. *Physics Reports*, 340:291–472, astro-ph/9912508.

Bartlett, J. G., Douspis, M., Blanchard, A., & Le Dour, M. (2000). An approximation to the likelihood function for band-power estimates of CMB anisotropies. *Astron. Astrophys. Suppl. Ser.*, 146:507–517, astro-ph/9903045.

Battye, R. A., Crittenden, R., & Weller, J. (2001). Cosmic concordance and the fine structure constant. *Phys. Rev.*, D63:043505, astro-ph/0008265.

Bean, R., Hansen, S. H., & Melchiorri, A. (2001). Early universe constraints on dark energy. *Phys. Rev.*, D64:103508, astro-ph/0104162.

Bean, R., Melchiorri, A., & Silk, J. (2003). Recombining WMAP: Beyond standard recombination. *Phys. Rev.*, D68:083501, astro-ph/0306357.

Becker, R. H. et al. (2001). Evidence for reionization at $z \sim 6$: Detection of a Gunn– Peterson trough in a $z = 6.28$ quasar. *Astron. J.*, 122:2850, astro-ph/0108097.

Bekenstein, J. D. (1982). Fine structure constant: Is it really a constant? *Phys. Rev.*, D25:1527–1539.

Bennett, C. L. et al. (1994). Cosmic temperature fluctuations from two years of COBE DMR observations. *Astrophys. J.*, 436:423–442, astro-ph/9401012.

Bennett, C. L. et al. (1996). 4-year COBE DMR cosmic microwave background observations: Maps and basic results. *Astrophys. J.*, 464:L1–L4, astro-ph/9601067.

Bennett, C. L. et al. (2003). First year Wilkinson Microwave Anisotropy Probe (WMAP) observations: Preliminary maps and basic results. *Astrophys. J. Suppl.*, 148:1, astro-ph/0302207.

Benoit, A. et al. (2003a). The cosmic microwave background anisotropy power spectrum measured by Archeops. *Astron. Astrophys.*, 399:L19–L23, astro-ph/0210305.

Benoit, A. et al. (2003b). Cosmological constraints from Archeops. *Astron. Astrophys.*, 399:L25–L30, astro-ph/0210306.

Blanchard, A., Douspis, M., Rowan-Robinson, M., & Sarkar, S. (2003). An alternative to the cosmological "concordance model". *Astron. Astrophys.*, 412:35–44, astro-ph/0304237.

Bond, J. R., Contaldi, C. R., & Pogosyan, D. (2003). Cosmic microwave background snapshots: pre–WMAP and post–WMAP. *Phil. Trans. Roy. Soc. Lond.*, A361:2435–2468, astro-ph/0310735.

Bond, J. R. & Efstathiou, G. (1984). Cosmic background radiation anisotropies in universes dominated by nonbaryonic dark matter. *Astrophys. J.*, 285:L45–L48.

Bond, J. R., Efstathiou, G., & Tegmark, M. (1997). Forecasting cosmic parameter errors from microwave background anisotropy experiments. *Mon. Not. Roy. Astron. Soc.*, 291:L33–L41, astro-ph/9702100.

Bond, J. R. et al. (2002). The Sunyaev-Zeldovich effect in CMB-calibrated theories applied to the Cosmic Background Imager anisotropy power at $\ell > 2000$. *Astrophys. J.*, in press, astro-ph/0205386.

Bond, J. R. & Jaffe, A. H. (1999). Constraining large scale structure theories with the cosmic background radiation. *Phil. Trans. Roy. Soc. London*, 357:57–75, astro-ph/9809043.

Bond, J. R., Jaffe, A. H., & Knox, L. E. (2000). Radical compression of cosmic microwave background data. *Astrophys. J.*, 533:19, astro-ph/9808264.

Bowen, R., Hansen, S. H., Melchiorri, A., Silk, J., & Trotta, R. (2002). The impact of an extra background of relativistic particles on the cosmological parameters derived from microwave background anisotropies. *Mon. Not. Roy. Astron. Soc.*, 334:760, astro-ph/0110636.

Box, G. E. & Tiao, G. C. (1973). *Bayesian inference in statistical analysis*. Addison-Wesley series in behavioral science: quantitative methods. Addison-Wesley, Reading, Massachusetts.

Braxmaier, C. et al. (2001). Proposed test of the time independence of the fundamental constants α and m_e/m_p using monolithic resonators. *Phys. Rev.*, D64:042001.

Bruscoli, M., Ferrara, A., & Scannapieco, E. (2002). How is the reionization epoch defined? *Mon. Not. Roy. Astron. Soc.*, 330:L43, astro-ph/0201094.

Bucher, M., Dunkley, J., Ferreira, P. G., Moodley, K., & Skordis, C. (2004). The initial conditions of the universe: how much isocurvature is allowed? astro-ph/0401417.

Bucher, M., Moodley, K., & Turok, N. (2000). The general primordial cosmic perturbation. *Phys. Rev.*, D62:083508, astro-ph/9904231.

Bucher, M., Moodley, K., & Turok, N. (2001). Constraining isocurvature perturbations with CMB polarization. *Phys. Rev. Lett.*, 87:191301, astro-ph/0012141.

Bucher, M., Moodley, K., & Turok, N. (2002). Characterising the primordial cosmic perturbations using MAP and PLANCK. *Phys. Rev.*, D66:023528, astro-ph/0007360.

Burles, S., Nollett, K. M., Truran, J. N., & Turner, M. S. (1999). Sharpening the predictions of Big-Bang Nucleosynthesis. *Phys. Rev. Lett.*, 82:4176–4179, astro-ph/9901157.

Burles, S., Nollett, K. M., & Turner, M. S. (2001). Big-Bang Nucleosynthesis predictions for precision cosmology. *Astrophys. J.*, 552:L1–L6, astro-ph/0010171.

Carlberg, R. G., Morris, S. L., Yee, H. K. C., & Ellingson, E. (1997). Redshift evolution of galaxy cluster densities. *Astrophys. J.*, 479:L19–L22, astro-ph/9612169.

Cen, R. (2003). The universe was reionized twice. *Astrophys. J.*, 591:12–37, astro-ph/0210473.

Christensen, N., Meyer, R., Knox, L., & Luey, B. (2001). Bayesian methods for cosmological parameter estimation from cosmic microwave background measurements. *Class. Quant. Grav.*, 18:2677, astro-ph/0103134.

Colless, M. et al. (2001). The 2dF Galaxy Redshift Survey: Spectra and redshifts. *Mon. Not. Roy. Astron. Soc.*, 328:1039, astro-ph/0106498.

Contaldi, C. R., Hoekstra, H., & Lewis, A. (2003). Joint CMB and weak lensing analysis: Physically motivated constraints on cosmological parameters. *Phys. Rev. Lett.*, 90:221303, astro-ph/0302435.

Cowles, M. K. & Carlin, B. P. (1996). Markov chain Monte Carlo convergence diagnostics: A comparative review. *Journal of the American Statistical Association*, 91(434):883–904.

Croft, R. A. C. et al. (2002). Towards a precise measurement of matter clustering: Lyman–alpha forest data at redshifts $2 - 4$. *Astrophys. J.*, 581:20–52, astro-ph/0012324.

Crotty, P., Garcia-Bellido, J., Lesgourgues, J., & Riazuelo, A. (2003a). Bounds on isocurvature perturbations from CMB and LSS data. *Phys. Rev. Lett.*, 91:171301, astro-ph/0306286.

Crotty, P., Lesgourgues, J., & Pastor, S. (2003b). Measuring the cosmological background of relativistic particles with WMAP. *Phys. Rev.*, D67:123005, astro-ph/0302337.

Cuoco, A. et al. (2003). Present status of primordial nucleosynthesis after WMAP: results from a new BBN code. astro-ph/0307213.

Cyburt, R. H., Fields, B. D., & Olive, K. A. (2003). Primordial nucleosynthesis in light of WMAP. *Phys. Lett.*, B567:227–234, astro-ph/0302431.

Damour, T. (2003a). String theory, cosmology and varying constants. *Astrophys. Space Sci.*, 283:445–456, gr-qc/0210059.

Damour, T. (2003b). Varying constants. Talk given at the 10th international workshop on Neutrino Telescopes, 11-14 March 2003 (Venice, Italy), gr-qc/0306023.

Damour, T. & Dyson, F. (1996). The Oklo bound on the time variation of the fine-structure constant revisited. *Nucl. Phys.*, B480:37–54, hep-ph/9606486.

Damour, T. & Nordtvedt, K. (1993). Tensor–scalar cosmological models and their relaxation toward general relativity. *Phys. Rev.*, D48:3436–3450.

de Bernardis, P. et al. (2002). Multiple peaks in the angular power spectrum of the cosmic microwave background: Significance and consequences for cosmology. *Astrophys. J.*, 564:559–566, astro-ph/0105296.

de Groot, S., van Leeuwen, W., & van Weert, C. (1980). *Relativistic kinetic theory.* North-Holland.

Di Bari, P. & Foot, R. (2001). Active-sterile neutrino oscillations and BBN + CMBR constraints. *Phys. Rev.*, D63:043008, hep-ph/0008258.

Dicus, D. A. et al. (1982). Primordial nucleosynthesis including radiative, Coulomb, and finite temperature corrections to weak rates. *Phys. Rev.*, D26:2694.

Djouadi, A., Drees, M., & Kneur, J. L. (2001). Constraints on the minimal supergravity model and prospects for SUSY particle production at future linear e+ e- colliders. *JHEP*, 08:055, hep-ph/0107316.

Dolgov, A. D. et al. (2002). Cosmological bounds on neutrino degeneracy improved by flavor oscillations. *Nucl. Phys.*, B632:363–382, hep-ph/0201287.

Dolgov, A. D. & Hansen, S. H. (2002). Massive sterile neutrinos as warm dark matter. *Astropart. Phys.*, 16:339–344, hep-ph/0009083.

Dolgov, A. D., Hansen, S. H., Pastor, S., & Semikoz, D. V. (1999). Unstable massive tau neutrinos and primordial nucleosynthesis. *Nucl. Phys.*, B548:385–407, hep-ph/9809598.

Dolgov, A. D., Hansen, S. H., Raffelt, G., & Semikoz, D. V. (2000). Heavy sterile neutrinos: Bounds from Big-Bang Nucleosynthesis and SN 1987A. *Nucl. Phys.*, B590:562–574, hep-ph/0008138.

Dolgov, A. D., Hansen, S. H., & Semikoz, D. V. (1997). Non-equilibrium corrections to the spectra of massless neutrinos in the early universe. *Nucl. Phys.*, B503:426–444, hep-ph/9703315.

Dolgov, A. D., Hansen, S. H., & Semikoz, D. V. (1998). Impact of massive tau neutrinos on primordial nucleosynthesis: Exact calculations. *Nucl. Phys.*, B524:621–638, hep-ph/9712284.

Doroshkevich, A. G., Naselsky, I. P., Naselsky, P. D., & Novikov, I. D. (2003). Ionization history of the cosmic plasma in the light of the recent CBI and future PLANCK data. *Astrophys. J.*, 586:709–717, astro-ph/0208114.

Durrer, R. (1990). Gauge invariant cosmological perturbation theory with seeds. *Phys. Rev.*, D42:2533.

Durrer, R. (1994). Gauge invariant cosmological perturbation theory: a general study and its application to the texture scenario of structure formation. *Fund. Cos. Phys.*, 15:209, astro-ph/9311041.

Durrer, R. (2001). The theory of CMB anisotropies. *J. Phys. Stud.*, 5:177–215, astro-ph/0109522.

Durrer, R., Gabrielli, A., Joyce, M., & Sylos Labini, F. (2003a). Bias and the power spectrum beyond the turn-over. *Astrophys. J.*, 585:L1–L4, astro-ph/0211653.

Durrer, R., Novosyadlyj, B., & Apunevych, S. (2003b). Acoustic peaks and dips in the CMB power spectrum: observational data and cosmological constraints. *Astrophys. J.*, 583:33–48, astro-ph/0111594.

Efstathiou, G. & Bond, J. R. (1999). Cosmic confusion: Degeneracies among cosmological parameters derived from measurements of microwave background anisotropies. *Mon. Not. Roy. Astron. Soc.*, 304:75–97, astro-ph/9807103.

Efstathiou, G., Bridle, S. L., Lasenby, A. N., Hobson, M. P., & Ellis, R. S. (1999). Constraints on Ω_Λ and Ω_m from distant type Ia supernovae and cosmic microwave background anisotropies. *Mon. Not. Roy. Astron. Soc.*, 303(3):L47–L52, astro-ph/9812226.

Einstein, A. (1917). *Sitzungsber K. Preuss. Akad. Wiss.*, pages 235–237.

Eisenstein, D. J., Hu, W., & Tegmark, M. (1998a). Cosmic complementarity: H_0 and Ω_m from combining CMB experiments and redshift surveys. *Astrophys. J.*, 504:L57–L60, astro-ph/9805239.

Eisenstein, D. J., Hu, W., & Tegmark, M. (1998b). Cosmic complementarity: Joint parameter estimation from CMB experiments and redshift surveys. *Astrophys. J.*, 518:2–23, astro-ph/9807130.

Ellis, J. R., Nanopoulos, D. V., & Olive, K. A. (2001). Combining the muon anomalous magnetic moment with other constraints on the CMSSM. *Phys. Lett.*, B508:65–73, hep-ph/0102331.

Enqvist, K., Kurki-Suonio, H., & Valiviita, J. (2000). Limits on isocurvature fluctuations from Boomerang and MAXIMA. *Phys. Rev.*, D62:103003, astro-ph/0006429.

Enqvist, K. & Sloth, M. S. (2002). Adiabatic CMB perturbations in pre Big-Bang string cosmology. *Nucl. Phys.*, B626:395–409, hep-ph/0109214.

Esposito, S., Mangano, G., Melchiorri, A., Miele, G., & Pisanti, O. (2001). Testing standard and degenerate Big-Bang Nucleosynthesis with BOOMERanG and MAXIMA-1. *Phys. Rev.*, D63:043004, astro-ph/0007419.

Esposito, S., Mangano, G., Miele, G., & Pisanti, O. (2000a). Big-Bang Nucleosynthesis: an accurate determination of light element yields. *Nucl. Phys.*, B568:421–444, astro-ph/9906232.

Esposito, S., Mangano, G., Miele, G., & Pisanti, O. (2000b). The standard and degenerate primordial nucleosynthesis versus recent experimental data. *JHEP*, 09:038, astro-ph/0005571.

Esposito, S., Miele, G., Pastor, S., Peloso, M., & Pisanti, O. (2000c). Non equilibrium spectra of degenerate relic neutrinos. *Nucl. Phys.*, B590:539–561, astro-ph/0005573.

Ferreira, P. G. & Joyce, M. (1997). Structure formation with a self-tuning scalar field. *Phys. Rev. Lett.*, 79:4740–4743, astro-ph/9707286.

Fields, B. & Sarkar, S. (2004). Big-Bang Nucleosynthesis. *Phys. Lett.*, B592(1). In: The Review of Particle Properties, Eidelman et al. (Particle Data Group), editors.

Filippenko, A. V. (2004). Evidence from type Ia supernovae for an accelerating universe and dark energy. In Freedman, W. L., editor, *Measuring and Modeling the Universe*, volume 2 of *Carnegie Observatories Astrophysics Series*, Cambridge. Cambridge Univ. Press, astro-ph/0307139.

Fixsen, D. J. et al. (1996). The cosmic microwave background spectrum from the full COBE FIRAS data set. *Ap. J.*, 473:576–587.

Fogli, G. L., Lisi, E., Marrone, A., Montanino, D., & Palazzo, A. (2001). Atmospheric, solar, and CHOOZ neutrinos: A global three generation analysis. Talk given at 36th Rencontres de Moriond on Electroweak Interactions and Unified Theories, Les Arcs, France, 10-17 Mar 2001, hep-ph/0104221.

Freedman, W. L. et al. (2001). Final results from the Hubble Space Telescope Key Project to measure the Hubble constant. *Astrophys. J.*, 553:47–72, astro-ph/0012376.

Frodesen, A., Skjeggestad, O., & Tofte, H. (1979). *Probability and statistics in particle physics*. Universitetsforlaget, Bergen-Oslo-Tromso, Norway.

Fujii, Y. (2002). Possible link between the changing fine structure constant and the accelerating universe via scalar-tensor theory. *Int. J. Mod. Phys.*, D11:1137–1148, astro-ph/0204069.

Fujii, Y. et al. (2000). The nuclear interaction at Oklo 2 billion years ago. *Nucl. Phys.*, B573:377–401, hep-ph/9809549.

Gamov, G. (1946). Expanding universe and the origin of elements. *Phys. Rev.*, 70:572–573.

Gawiser, E. (2001). Interpreting CMB anisotropy observations: Trying to tell the truth with statistics. In Wheeler, J. & Martel, N., editors, *Austin 2000, Relativistic Astrophysics*, pages 202–207. astro-ph/0105010. Proceedings of 20th Texas Symposium on Relativistic Astrophysics, Austin, Texas, 11-15 Dec 2000. AIP.

Gelman, A. & Rubin, D. (1992a). A single series from the Gibbs sampler provides a false sense of security. In Bernardo, J. et al., editors, *Bayesian Statistics*, volume 4, pages 625–632. Oxford University Press.

Gelman, A. & Rubin, D. (1992b). Inference from iterative simulation using multiple sequences (with discussion). *Statistical Science*, 7:457–511.

Gilks, W., Richardson, S., & Spiegelhalter, D., editors (1996). *Markov chain Monte Carlo in practice*. Chapman & Hall, London.

Goldstein, J. H. et al. (2003). Estimates of cosmological parameters using the CMB angular power spectrum of ACBAR. *Astrophys. J.*, 599:773–785, astro-ph/0212517.

Gonzalez-Garcia, M. C., Maltoni, M., Pena-Garay, C., & Valle, J. W. F. (2001). Global three-neutrino oscillation analysis of neutrino data. *Phys. Rev.*, D63:033005, hep-ph/0009350.

Goobar, A. & Perlmutter, S. (1995). Feasibility of measuring the cosmological constant Lambda and mass density Omega using type Ia supernovae. *Astrophys. J.*, 450:14, astro-ph/9505022.

Gordon, C. & Lewis, A. (2003). Observational constraints on the curvaton model of inflation. *Phys. Rev.*, D67:123513, astro-ph/0212248.

Gordon, C. & Malik, K. A. (2004). WMAP, neutrino degeneracy and non-Gaussianity constraints on isocurvature perturbations in the curvaton model of inflation. *Phys. Rev.*, D69:063508, astro-ph/0311102.

Gradshteyn, I. & Ryzhik, I. (1965). *Tables of Integrals, Series and Products*. Academic Press, NY, 1st edition.

Haiman, Z. (2004). The first nonlinear structures and the reionization history of the universe. In Ho, L., editor, *Coevolution of Black Holes and Galaxies*, volume 1 of *Carnegie Observatories Astrophysics Symposium*, Cambridge. Cambridge University Press, astro-ph/0304131.

Haiman, Z. & Holder, G. P. (2003). The reionization history at high redshifts I: Physical models and new constraints from CMB polarization. *Astrophys. J.*, 595:1–12, astro-ph/0302403.

Halverson, N. W. et al. (2002). DASI first results: A measurement of the cosmic microwave background angular power spectrum. *Astrophys. J.*, 568:38–45, astro-ph/0104489.

Hamilton, A. J. S. & Tegmark, M. (2002). The real space power spectrum of the PSCz survey from 0.01 to 300 h/Mpc. *Mon. Not. Roy. Astron. Soc.*, 330:506, astro-ph/0008392.

Hamuy, M. et al. (1996). The morphology of type Ia supernovae light curves. *Astronomical Journal*, 112(2391), astro-ph/9609063.

Hanany, S. et al. (2000). MAXIMA-1: A measurement of the cosmic microwave background anisotropy on angular scales of 10 arcminutes to 5 degrees. *Astrophys. J.*, 545:L5, astro-ph/0005123.

Hannestad, S. (1999). Possible constraints on the time variation of the fine structure constant from cosmic microwave background data. *Phys. Rev.*, D60:023515, astro-ph/9810102.

Hannestad, S. (2000). New constraints on neutrino physics from Boomerang data. *Phys. Rev. Lett.*, 85:4203–4206, astro-ph/0005018.

Hannestad, S. (2001). New CMBR data and the cosmic neutrino background. *Phys. Rev.*, D64:083002, astro-ph/0105220.

Hannestad, S. (2003). Neutrino masses and the number of neutrino species from WMAP and 2dFGRS. *JCAP*, 0305:004, astro-ph/0303076.

Hansen, S. H. & Haiman, Z. (2004). Do we need stars to reionize the universe at high redshifts? Early reionization by decaying heavy sterile neutrinos. *Astrophys. J.*, 600:26–31, astro-ph/0305126.

Hansen, S. H., Mangano, G., Melchiorri, A., Miele, G., & Pisanti, O. (2002). Constraining neutrino physics with BBN and CMBR. *Phys. Rev.*, D65:023511, astro-ph/0105385.

Hansen, S. H. & Villante, F. L. (2000). Decaying neutrino and a high cosmological baryon density. *Phys. Lett.*, B486:1–5, astro-ph/0005114.

Harrison, E. R. (1970). Fluctuations at the threshold of classical cosmology. *Phys. Rev.*, D1:2726–2730.

Hastings, W. (1970). Monte Carlo sampling methods using Markov chains and their applications. *Biometrika*, 57:97–109.

Heckler, A. F. (1994). Astrophysical applications of quantum corrections to the equation of state of a plasma. *Phys. Rev.*, D49:611–617.

Heuser, H. (1993). *Lehrbuch der Analysis. Teil 2.* B.G. Teubner Stuttgart, 8th edition.

Hinshaw, G. et al. (2003a). First year Wilkinson Microwave Anisotropy Probe (WMAP) observations: Angular power spectrum. *Astrophys. J. Suppl.*, 148:135, astro-ph/0302217.

Hinshaw, G. et al. (2003b). First year Wilkinson Microwave Anisotropy Probe (WMAP) observations: Data processing methods and systematic errors limits. *Astrophys. J. Suppl.*, 148:63, astro-ph/0302222.

Holder, G., Haiman, Z., Kaplinghat, M., & Knox, L. (2003). The reionization history at high redshifts II: Estimating the optical depth to Thomson scattering from CMB polarization. *Astrophys. J.*, 595:13–18, astro-ph/0302404.

Hu, W. & Dodelson, S. (2002). Cosmic microwave background anisotropies. *Ann. Rev. Astron. Astrophys.*, 40:171, astro-ph/0110414.

Hu, W., Eisenstein, D. J., Tegmark, M., & White, M. J. (1999). Observationally determining the properties of dark matter. *Phys. Rev.*, D59:023512, astro-ph/9806362.

Hu, W. & Holder, G. P. (2003). Model-independent reionization observables in the CMB. *Phys. Rev.*, D68:023001, astro-ph/0303400.

Hu, W., Scott, D., Sugiyama, N., & White, M. J. (1995). The Effect of physical assumptions on the calculation of microwave background anisotropies. *Phys. Rev.*, D52:5498–5515, astro-ph/9505043.

Hu, W. & Sugiyama, N. (1995a). Anisotropies in the cosmic microwave background: An analytic approach. *Astrophys. J.*, 444:489–506.

Hu, W. & Sugiyama, N. (1995b). Toward understanding CMB anisotropies and their implications. *Phys. Rev.*, D51:2599–2630, astro-ph/9411008.

Hu, W. & Sugiyama, N. (1996). Small scale cosmological perturbations: An analytic approach. *Astrophys. J.*, 471:542–570, astro-ph/9510117.

Hu, W., Sugiyama, N., & Silk, J. (1997). The physics of microwave background anisotropies. *Nature*, 386:37–43, astro-ph/9504057.

Hu, W. & White, M. (1997). The damping tail of CMB anisotropies. *Astrophys. J.*, 479:568, astro-ph/9609079.

Ivanchik, A., Petitjean, P., Rodriguez, E., & Varshalovich, D. (2003). Does the proton-to-electron mass ratio vary in the course of cosmological evolution? *Astrophys. Space Sci.*, 283:583–588, astro-ph/0210299.

Ivanchik, A. V., Rodriguez, E., Petitjean, P., & Varshalovich, D. A. (2002). Do the fundamental constants vary in the course of the cosmological evolution? *Sov. Astron. Lett.*, 28:423, astro-ph/0112323.

Izotov, Y. I. et al. (1999). Helium abundance in the most metal-deficient blue compact galaxies: I Zw 18 and SBS 0335-052. *Astrophys. J.*, 527:757, astro-ph/9907228.

Izotov, Y. I. & Thuan, T. X. (1998). The primordial abundance of ^{4}He revisited. *Astrophys. J.*, 500:188.

Jackson, J. (1975). *Classical Electrodynamics*. John Wiley & Sons, 2nd edition.

Jaffe, A. H., Bond, J. R., Ferreira, P. G., & Knox, L. E. (2003). CMB likelihood functions for beginners and experts. In Melchiorri, F., editor, *AIP Conf. Proc.*, volume 476, page 249. astro-ph/0306506.

Jaynes, E. (2003). *Probability theory. The logic of science*. Cambridge University Press, Cambridge, UK.

Jimenez, R., Verde, L., Peiris, H., & Kosowsky, A. (2004). Fast cosmological parameter estimation from microwave background temperature and polarization power spectra. *Phys. Rev.*, D70:023005, astro-ph/0404237.

Jones, B. J. (1997). Statistics and random functions in astrophysics. In Lineweaver et al. (1997), pages 67–110. Proceedings of the NATO Advanced Study Institute on the cosmological background radiation, Strasbourg, France, May 27-June 7, 1996.

Kaiser, N. (1984). On the spatial correlations of Abell clusters. *Astrophys. J.*, 284:L9–L12.

Kaiser, N. (1987). Clustering in real space and in redshift space. *Mon. Not. Roy. Astron. Soc.*, 227:1–27.

Kaiser, N. & Squires, G. (1993). Mapping the dark matter with weak gravitational lensing. *Astrophys. J.*, 404:441–450.

Kamionkowski, M., Kosowsky, A., & Stebbins, A. (1997). Statistics of cosmic microwave background polarization. *Phys. Rev.*, D55:7368–7388, astro-ph/9611125.

Kang, H.-S. & Steigman, G. (1992). Cosmological constraints on neutrino degeneracy. *Nucl. Phys.*, B372:494–520.

Kaplinghat, M. et al. (2003a). Probing the reionization history of the universe using the cosmic microwave background polarization. *Astrophys. J.*, 583:24–32, astro-ph/0207591.

Kaplinghat, M., Knox, L., & Song, Y.-S. (2003b). Determining neutrino mass from the CMB alone. *Phys. Rev. Lett.*, 91:241301, astro-ph/0303344.

Kaplinghat, M., Scherrer, R. J., & Turner, M. S. (1999). Constraining variations in the fine-structure constant with the cosmic microwave background. *Phys. Rev.*, D60:023516, astro-ph/9810133.

Kaplinghat, M. & Turner, M. S. (2001). Precision cosmology and the density of baryons in the universe. *Phys. Rev. Lett.*, 86:385, astro-ph/0007454.

Kendall, M. & Stuart, A. (1977). *The advanced theory of statistics*. Griffin & Co, London & High Wycombe, fourth edition. 3 volumes.

Kinney, W. H., Kolb, E. W., Melchiorri, A., & Riotto, A. (2004). WMAPping inflationary physics. *Phys. Rev.*, D69:103516, hep-ph/0305130.

Kinney, W. H. & Riotto, A. (1999). Measuring the cosmological lepton asymmetry through the CMB, anisotropy. *Phys. Rev. Lett.*, 83:3366–3369, hep-ph/9903459.

Kirilova, D. P. (2003). Overproduction of primordial helium-4 in the presence of neutrino oscillations. *Astropart. Phys.*, 19:409–417, astro-ph/0109105.

Kirkman, D., Tytler, D., Suzuki, N., O'Meara, J. M., & Lubin, D. (2003). The cosmological baryon density from the deuterium to, hydrogen ratio towards QSO absorption systems: D/H towards Q1243+3047. *Astrophys. J. Suppl.*, 149:1, astro-ph/0302006.

Kneller, J. P., Scherrer, R. J., Steigman, G., & Walker, T. P. (2001). When Does CMB + BBN = New Physics? *Phys. Rev.*, D64:123506, astro-ph/0101386.

Knox, L. (1995). Determination of inflationary observables by cosmic microwave background anisotropy experiments. *Phys. Rev.*, D52:4307–4318, astro-ph/9504054.

Knox, L. & Page, L. (2000). Characterizing the peak in the CMB angular power spectrum. *Phys. Rev. Lett.*, 85:1366–1369, astro-ph/0002162.

Kochanek, C. S. & Schechter, P. L. (2004). The Hubble constant from gravitational lens time delays. In Freedman, W. L., editor, *Measuring and Modeling the Universe*, volume 2 of *Carnegie Observatories Astrophysics Series*, Cambridge. Cambridge Univ. Press, astro-ph/0306040.

Kodama, H. & Sasaki, M. (1984). Cosmological perturbation theory. *Prog. Theor. Phys. Suppl.*, 78:1.

Kogut, A. et al. (2003). Wilkinson Microwave Anisotropy Probe (WMAP) first year observations: TE polarization. *Astrophys. J. Suppl.*, 148:161, astro-ph/0302213.

Kolb, E. W. & Turner, M. S. (1990). *The early universe*. Frontiers in Physics. Addison-Wesley Publishing Company, 1st edition.

Kolb, E. W., Turner, M. S., Chakravorty, A., & Schramm, D. N. (1991). Constraints from primordial nucleosynthesis on the mass of the tau neutrino. *Phys. Rev. Lett.*, 67:533–536.

Kosowsky, A. (1996). Cosmic microwave background polarization. *Ann. Phys.*, 246:49–85, astro-ph/9501045.

Kosowsky, A. (2002). The cosmic microwave background. In Bonometto, S., Gorini, V., & Moschella, U., editors, *Modern Cosmology*, pages 219–263, Bristol. IOP Publishing, astro-ph/0102402.

Kosowsky, A., Kamionkowski, M., Jungman, G., & Spergel, D. N. (1996). Determining cosmological parameters from the microwave background. *Nucl. Phys. Proc. Suppl.*, 51B:49–53, astro-ph/9605147.

Kosowsky, A., Milosavljevic, M., & Jimenez, R. (2002). Efficient cosmological parameter estimation from microwave background anisotropies. *Phys. Rev.*, D66:063007, astro-ph/0206014.

Kovac, J. et al. (2002). Detection of polarization in the cosmic microwave background using DASI. *Nature*, 420:772–787, astro-ph/0209478.

Kuo, C.-l. et al. (2004). High resolution observations of the CMB power spectrum with ACBAR. *Astrophys. J.*, 600:32–51, astro-ph/0212289.

Lange, A. E. et al. (2001). Cosmological parameters from the first results of BOOMERANG. *Phys. Rev.*, D63:042001, astro-ph/0005004.

Langlois, D. (1999). Correlated adiabatic and isocurvature perturbations from double inflation. *Phys. Rev.*, D59:123512.

Langlois, D. & Riazuelo, A. (2000). Correlated mixtures of adiabatic and isocurvature cosmological perturbations. *Phys. Rev.*, D62:043504, astro-ph/9912497.

Lee, A. T. et al. (2001). A high spatial resolution analysis of the MAXIMA-1 cosmic microwave background anisotropy data. *Astrophys. J.*, 561:L1–L6, astro-ph/0104459.

Lesgourgues, J. & Pastor, S. (1999). Cosmological implications of a relic neutrino asymmetry. *Phys. Rev.*, D60:103521, hep-ph/9904411.

Lewis, A. & Bridle, S. (2002). Cosmological parameters from CMB and other data: a Monte-Carlo approach. *Phys. Rev.*, D66:103511, astro-ph/0205436.

Lewis, A., Challinor, A., & Lasenby, A. (2000). Efficient computation of CMB anisotropies in closed FRW models. *Astrophys. J.*, 538:473–476, astro-ph/9911177.

Libarskii, I. & Sunyaev, R. (1983). The spectral features in the microwave background spectrum due to energy release in the early universe. *Astron. Astrophys.*, 123:171.

Liddle, A. R. (2004). How many cosmological parameters? *Mon. Not. Roy. Astron. Soc.*, astro-ph/0401198. To appear.

Liddle, A. R. & Lyth, D. H. (2000). *Cosmological inflation and large-scale structure*. Cambridge University Press, Cambridge, UK.

Lineweaver, C., Bartlett, J., Blanchard, A., Signore, M., & Silk, J., editors (1997). *The cosmic microwave background*, volume 502 of *NATO ASI Series; Series C: Mathematical and Physical Sciences*. Kluver Academic. Proceedings of the NATO Advanced Study Institute on the cosmological background radiation, Strasbourg, France, May 27-June 7, 1996.

Lisi, E., Sarkar, S., & Villante, F. L. (1999). The Big-Bang Nucleosynthesis limit on N_ν. *Phys. Rev.*, D59:123520, hep-ph/9901404.

Lopez, R. E., Dodelson, S., Heckler, A., & Turner, M. S. (1999). Precision detection of the cosmic neutrino background. *Phys. Rev. Lett.*, 82:3952–3955, astro-ph/9803095.

Lopez, R. E. & Turner, M. S. (1999). An accurate calculation of the Big-Bang prediction for the abundance of primordial helium. *Phys. Rev.*, D59:103502, astro-ph/9807279.

Luridiana, V., Peimbert, A., Peimbert, M., & Cervino, M. (2003). The effect of collisional enhancement of Balmer lines on the determination of the primordial helium abundance. *Astrophys. J.*, 592:846–865, astro-ph/0304152.

Lyth, D. H., Ungarelli, C., & Wands, D. (2003). The primordial density perturbation in the curvaton scenario. *Phys. Rev.*, D67:023503, astro-ph/0208055.

Lyth, D. H. & Wands, D. (2002). Generating the curvature perturbation without an inflaton. *Phys. Lett.*, B524:5–14, hep-ph/0110002.

Lyth, D. H. & Wands, D. (2003). The CDM isocurvature perturbation in the curvaton scenario. *Phys. Rev.*, D68:103516, astro-ph/0306500.

Ma, C. & Bertschinger, E. (1995). Cosmological perturbation theory in the synchronous and conformal newtonian gauges. *Astrophys. J.*, 455:7–25.

MacKay, D. J. (2003). *Information theory, inference, and learning algorithms*. Cambridge University Press, Cambridge.

Malik, K. A., Wands, D., & Ungarelli, C. (2003). Large-scale curvature and entropy perturbations for multiple interacting fluids. *Phys. Rev.*, D67:063516, astro-ph/0211602.

Mandelbaum, R., McDonald, P., Seljak, U., & Cen, R. (2003). Precision cosmology from the Lyman-alpha forest: Power spectrum and bispectrum. *Mon. Not. Roy. Astron. Soc.*, 344:776, astro-ph/0302112.

Mangano, G., Miele, G., Pastor, S., & Peloso, M. (2002). A precision calculation of the effective number of cosmological neutrinos. *Phys. Lett.*, B534:8–16, astro-ph/0111408.

Martins, C. J. A. P. (2002). Cosmology with varying constants. *Phil. Trans. Roy. Soc. Lond.*, A360:2681–2696, astro-ph/0205504.

Martins, C. J. A. P., Melchiorri, A., Rocha, G., Trotta, R., Avelino, P., & Viana, P. (2004). WMAP constraints on varying α and the promise of reionization. *Phys. Lett. B*, 585:29, astro-ph/0302295.

Martins, C. J. A. P., Melchiorri, A., Trotta, R., Bean, R., Rocha, G., Avelino, P., & Viana, P. (2002). Measuring α in the early universe: CMB temperature, large-scale structure and Fisher matrix analysis. *Phys. Rev.*, D66:023505, astro-ph/0203149.

Mason, B. S. et al. (2003). The anisotropy of the microwave background to $\ell = 3500$: Deep Field observations with the Cosmic Background Imager. *Astrophys. J.*, 591:540–555, astro-ph/0205384.

Melchiorri, A. et al. (2000). A measurement of Omega from the North American test flight of BOOMERANG. *Astrophys. J.*, 536:L63–L66, astro-ph/9911445.

Melchiorri, A. & Griffiths, L. M. (2001). From Anisotropy to Omega. *New Astron. Rev.*, 45:321–328, astro-ph/0011147.

Mengersen, K., Knight, S., & Robert, C. (1999). MCMC: how do we know when to stop? In *Bulletin of the International Statistical Institute*, volume LVIII, 52th session.

Meszaros, A. (2002). On the reality of the accelerating universe. *Astrophys. J.*, 580:12–15, astro-ph/0207558.

Meszaros, P. (1974). The behaviour of point masses in an expanding cosmological substratum. *Astron. Astrophys.*, 37:225–228.

Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H., & Teller, E. (1953). Equation of state calculations by fast computing machines. *J. Chem. Phys.*, 21:1087–1092.

Miller, A. D. et al. (1999). A measurement of the angular power spectrum of the CMB from $\ell = 100$ to 400. *Astrophys. J.*, 524:L1–L4, astro-ph/9906421.

Moroi, T. & Takahashi, T. (2001). Effects of cosmological moduli fields on cosmic microwave background. *Phys. Lett.*, B522:215–221, hep-ph/0110096.

Mota, D. F. & Barrow, J. D. (2004). Varying alpha in a more realistic universe. *Phys. Lett.*, B581:141–146, astro-ph/0306047.

Mukhanov, V. F., Feldman, H. A., & Brandenberger, R. H. (1992). Theory of cosmological perturbations. Part 1. Classical perturbations. Part 2. Quantum theory of perturbations. Part 3. Extensions. *Phys. Rept.*, 215:203–333.

Murphy, M. T. et al. (2001a). Improved constraints on possible variation of physical constants from H I 21cm and molecular QSO absorption lines. *Mon. Not. Roy. Astron. Soc.*, 327:1244, astro-ph/0101519.

Murphy, M. T. et al. (2001b). Possible evidence for a variable fine structure constant from QSO absorption lines: motivations, analysis and results. *Mon. Not. Roy. Astron. Soc.*, 327:1208, astro-ph/0012419.

Murphy, M. T., Webb, J. K., Flambaum, V. V., & Curran, S. J. (2003). Does the fine structure constant vary? A detailed investigation into systematic effects. *Astrophys. Space Sci.*, 283:577, astro-ph/0210532.

Murphy, M. T., Webb, J. K., Flambaum, V. V., Prochaska, J. X., & Wolfe, A. M. (2001c). Further constraints on variation of the fine structure constant from alkali doublet QSO absorption lines. *Mon. Not. Roy. Astron. Soc.*, 327:1237, astro-ph/0012421.

Netterfield, C. B., Devlin, M. J., Jarosik, N., Page, L., & Wollack, E. J. (1997). A measurement of the angular power spectrum of the anisotropy in the cosmic microwave background. *Astrophys. J.*, 474:47–54, astro-ph/9601197.

Netterfield, C. B. et al. (2002). A measurement by BOOMERANG of multiple peaks in the angular power spectrum of the cosmic microwave background. *Astrophys. J.*, 571:604–614, astro-ph/0104460.

Olive, K. A. & Pospelov, M. (2002). Evolution of the fine structure constant driven by dark matter and the cosmological constant. *Phys. Rev.*, D65:085044, hep-ph/0110377.

Olive, K. A., Skillman, E., & Steigman, G. (1997). The primordial abundance of he4: An update. *Astrophys. J.*, 483:788, astro-ph/9611166.

Olive, K. A., Steigman, G., & Walker, T. P. (2000). Primordial nucleosynthesis: Theory and observations. *Phys. Rept.*, 333:389–407, astro-ph/9905320.

Padin, S. et al. (2001). First intrinsic anisotropy observations with the Cosmic Background Imager. *Astrophys. J.*, 549:L1–L5, astro-ph/0012211.

Padmanabhan, T. (1993). *Structure formation in the universe*. Cambridge University Press, Cambrige, UK.

Page, L. et al. (2003). First year Wilkinson Microwave Anisotropy Probe (WMAP) observations: Interpretation of the TT and TE angular power spectrum peaks. *Astrophys. J. Suppl.*, 148:233, astro-ph/0302220.

Partridge, R. B. (1995). *3 K: the cosmic microwave background radiation*. Number 25 in Cambridge astrophysics series. Cambridge University Press, Cambridge.

Pearson, T. J. et al. (2003). The anisotropy of the microwave background to $\ell = 3500$: Mosaic observations with the Cosmic Background Imager. *Astrophys. J.*, 591:556–574, astro-ph/0205388.

Peebles, P. J. E. (1980). *Large scale structure of the universe*. Princeton University Press, Princeton, USA.

Peimbert, A., Peimbert, M., & Luridiana, V. (2002). Temperature bias and the primordial Helium abundance determination. *Astrophys. J.*, 565:668, astro-ph/0107189.

Peiris, H. V. et al. (2003). First year Wilkinson Microwave Anisotropy Probe (WMAP) observations: Implications for inflation. *Astrophys. J. Suppl.*, 148:213, astro-ph/0302225.

Penzias, A. A. & Wilson, R. W. (1965). A measurement of excess antenna temperature at 4080-Mc/s. *Astrophys. J.*, 142:419.

Perlmutter, S. et al. (1999). Measurements of Ω and Λ from 42 high-redshift supernovae. *Astrophys. J.*, 517:565–586, astro-ph/9812133.

Pettini, M. & Bowen, D. V. (2001). A new measurement of the primordial abundance of Deuterium: Toward convergence with the baryon density from the CMB? *Astrophys. J.*, 560:41–48, astro-ph/0104474.

Pierpaoli, E. (2003). Constraints on the cosmic neutrino background. *Mon. Not. Roy. Astron. Soc.*, 342:L63, astro-ph/0302465.

Pierpaoli, E., Garcia-Bellido, J., & Borgani, S. (1999). Microwave background anisotropies and large scale structure constraints on isocurvature modes in a two-field model of inflation. *JHEP*, 10:015, hep-ph/9909420.

Planck Website (2004). `http://astro.estec.esa.nl/Planck/`.

Polchinski, J. (1998). *String theory. Vol. 1: An introduction to the bosonic string. Vol. 2: Superstring theory and beyond*. Cambridge University Press, Cambridge, UK.

Press, W., Teukolsky, S., W.T., V., & Flannery, B. (1992). *Numerical Recipes in Fortran. The art of scientific computing*. Cambridge University Press, Cambridge, UK, 2nd edition.

Pryke, C. et al. (2002). Cosmological parameter extraction from the first season of observations with DASI. *Astrophys. J.*, 568:46–51, astro-ph/0104490.

RADPack Website (2001). `http://bubba.ucdavis.edu/~knox/radpack.html`.

Raftery, A. E. & Lewis, S. M. (1996). The number of iterations, convergence diagnostics and generic Metropolis algorithms. In Gilks, W., Spiegelhalter, D., & Richardson, S., editors, *Practical Markov Chain Monte Carlo*. Chapman & Hall, London.

Ratra, B. & Peebles, P. J. E. (1988). Cosmological consequences of a rolling homogeneous scalar field. *Phys. Rev.*, D37:3406.

Readhead, A. C. S. et al. (2004). Extended Mosaic observations with the Cosmic Background Imager. *Astrophys. J.*, 609:498–512, astro-ph/0402359.

Riazuelo, A., Vernizzi, F., Steer, D., & Durrer, R. (2002). Gauge invariant cosmological perturbation theory for braneworlds. hep-th/0205220.

Riess, A. G. et al. (1998). Observational evidence from supernovae for an accelerating universe and a cosmological constant. *Astron. J.*, 116:1009–1038, astro-ph/9805201.

Riess, A. G. et al. (2004). Type Ia supernova discoveries at $z > 1$ from the Hubble Space Telescope: Evidence for past deceleration and constraints on dark energy evolution. *Astrophys. J.*, 607:665–687, astro-ph/0402512.

Riess, A. G., Press, W. H., & Kirshner, R. P. (1996). A precise distance indicator: Type Ia supernova multicolor light curve shapes. *Astrophys. J.*, 473:88, astro-ph/9604143.

Rocha, G., Trotta, R., Martins, C., Melchiorri, A., Avelino, P., Bean, R., & Viana, P. (2004). Measuring α in the early Universe: CMB polarization, reionization and the Fisher matrix analysis. *Mon. Not. Roy. Astron. Soc.*, 352:20–38, astro-ph/0309211.

Rubino-Martin, J. A. et al. (2003). First results from the Very Small Array. IV: Cosmological parameter estimation. *Mon. Not. Roy. Astron. Soc.*, 341:1084, astro-ph/0205367.

Ryan, S. G., Beers, T. C., Olive, K. A., Fields, B. D., & Norris, J. E. (2000). Primordial Lithium and Big-Bang Nucleosynthesis. *Astrophys. J.*, 530:L57, astro-ph/9905211.

Sahni, V. (2004). Dark matter and dark energy. Lectures given at 2nd Aegean Summer School on the Early Universe, Ermoupoli, Island of Syros, Greece, 22-30 Sep 2003, astro-ph/0403324.

Salaris, M. & Weiss, A. (2001). Atomic diffusion in metal-poor stars: II. Predictions for the Spite plateau. *Astron. Astrophys.*, 376:955–965, astro-ph/0104406.

Sandvik, H. B., Barrow, J. D., & Magueijo, J. (2002). A simple varying-alpha cosmology. *Phys. Rev. Lett.*, 88:031302, astro-ph/0107512.

Sandvik, H. B., Tegmark, M., Wang, X., & Zaldarriaga, M. (2004). CMBfit: Rapid WMAP likelihood calculations with normal parameters. *Phys. Rev.*, D69:063005, astro-ph/0311544.

Santiago, D. I., Kalligas, D., & Wagoner, R. V. (1998). Scalar-tensor cosmologies and their late time evolution. *Phys. Rev.*, D58:124005, gr-qc/9805044.

Scott, P. F. et al. (2003). First results from the Very Small Array – III. The CMB power spectrum. *Mon. Not. Roy. Astron. Soc.*, 341:1076, astro-ph/0205380.

Seager, S., Sasselov, D. D., & Scott, D. (1999). A new calculation of the recombination epoch. *Astrophys. J.*, 523:L1, astro-ph/9909275.

Seager, S., Sasselov, D. D., & Scott, D. (2000). How exactly did the Universe become neutral? *Astrophys. J. Suppl.*, 128:407–430, astro-ph/9912182.

Seljak, U., McDonald, P., & Makarov, A. (2003a). Cosmological constraints from the CMB and Ly-alpha forest revisited. *Mon. Not. Roy. Astron. Soc.*, 342:L79, astro-ph/0302571.

Seljak, U., Sugiyama, N., White, M. J., & Zaldarriaga, M. (2003b). A comparison of cosmological Boltzmann codes: are we ready for high precision cosmology? *Phys. Rev.*, D68:083507, astro-ph/0306052.

Seljak, U. & Zaldarriaga, M. (1996). A line of sight approach to cosmic microwave background anisotropies. *Astrophys. J.*, 469:437–444, astro-ph/9603033.

Shi, X.-d., Fuller, G. M., & Abazajian, K. (1999). Neutrino-mixing-generated lepton asymmetry and the primordial He-4 abundance. *Phys. Rev.*, D60:063002, astro-ph/9905259.

Sievers, J. L. et al. (2003). Cosmological parameters from Cosmic Background Imager observations and comparisons with BOOMERANG, DASI, and MAXIMA. *Astrophys. J.*, 591:599–622, astro-ph/0205387.

Sigurdson, K., Kurylov, A., & Kamionkowski, M. (2003). Spatial variation of the fine-structure parameter and the cosmic microwave background. *Phys. Rev.*, D68:103509, astro-ph/0306372.

Silk, J. (1968). Cosmic black body radiation and galaxy formation. *Astrophys. J.*, 151:459–471.

Skordis, C. & Albrecht, A. (2002). Planck-scale quintessence and the physics of structure formation. *Phys. Rev.*, D66:043523, astro-ph/0012195.

Skordis, C. & Silk, J. (2004). A new method for measuring the CMB temperature quadrupole with an accuracy better than cosmic variance. astro-ph/0402474.

Slosar, A. & Hobson, M. (2003). An improved Markov-chain Monte Carlo sampler for the estimation of cosmological parameters from CMB data. *Mon. Not. Roy. Astron. Soc.*, in press, astro-ph/0307219.

Smoot, G. F. et al. (1992). Structure in the COBE DMR first year maps. *Astrophys. J.*, 396:L1–L5.

Spergel, D. N. et al. (2003). First year Wilkinson Microwave Anisotropy Probe (WMAP) observations: Determination of cosmological parameters. *Astrophys. J. Suppl.*, 148:175, astro-ph/0302209.

Steigman, G. (2001). Precision neutrino counting. astro-ph/0108148.

Stewart, J. M. & Walker, M. (1974). Perturbations of space-times in general relativity. *Proc. R. Soc. London*, A341(1624):49–74.

Stompor, R., Banday, A. J., & Gorski, Krzysztof, M. (1996). Flat dark matter dominated models with hybrid adiabatic plus isocurvature initial conditions. *Astrophys. J.*, 463:8, astro-ph/9511087.

Stompor, R. et al. (2001). Cosmological implications of the MAXIMA-I high resolution Cosmic Microwave Background anisotropy measurement. *Astrophys. J.*, 561:L7–L10, astro-ph/0105062.

Straumann, N. (2003). On the cosmological constant problems and the astronomical evidence for a homogeneous energy density with negative pressure. In Duplantier, B. & Rivasseau, V., editors, *Vacuum Energy–Renormalization*, volume 30 of *Progress in Mathematical Physics*, pages 7–51, Basel. Birkhäuser, astro-ph/0203330.

Straumann, N. (2004). *General Relativity with Applications to Astrophysics*. Springer, Berlin.

Taylor, A. C. et al. (2003). First results from the Very Small Array – II. Observations of the CMB. *Mon. Not. Roy. Astron. Soc.*, 341:1066, astro-ph/0205381.

Tegmark, M., Eisenstein, D. J., Hu, W., & de Oliveira-Costa, A. (2000). Foregrounds and forecasts for the cosmic microwave background. *Astrophys. J.*, 530:133–165, astro-ph/9905257.

Tegmark, M. et al. (2004a). The 3D power spectrum of galaxies from the SDSS. *Astrophys. J.*, 606:702–740, astro-ph/0310725.

Tegmark, M. et al. (2004b). Cosmological parameters from SDSS and WMAP. *Phys. Rev.*, D69:103501, astro-ph/0310723.

Tegmark, M. & Hamilton, A. J. S. (1997). Uncorrelated measurements of the CMB power spectrum. In Olinto, A. et al., editors, *18th Texas Symposium on relativistic astrophysics and cosmology*, page 270. World Scientific, Singapore, astro-ph/9702019.

Tegmark, M., Hamilton, A. J. S., & Xu, Y. (2002). The power spectrum of galaxies in the 2dF 100k redshift survey. *Mon. Not. Roy. Astron. Soc.*, 335:887–908, astro-ph/0111575.

Tegmark, M. & Silk, J. (1995). Reionization in an open CDM universe: implications for cosmic microwave background fluctuations. *Astrophys. J.*, 441:458–464, astro-ph/9405042.

Tegmark, M., Taylor, A., & Heavens, A. (1997). Karhunen-Loeve eigenvalue problems in cosmology: how should we tackle large data sets? *Astrophys. J.*, 480:22, astro-ph/9603021.

Tegmark, M. & Zaldarriaga, M. (2000). New CMB constraints on the cosmic matter budget: trouble for nucleosynthesis? *Phys. Rev. Lett.*, 85:2240, astro-ph/0004393.

Tegmark, M. & Zaldarriaga, M. (2002). Separating the early universe from the late universe: cosmological parameter estimation beyond the black box. *Phys. Rev.*, D66:103508, astro-ph/0207047.

Tegmark, M., Zaldarriaga, M., & Hamilton, A. J. S. (2001). Towards a refined cosmic concordance model: joint 11-parameter constraints from CMB and large-scale structure. *Phys. Rev.*, D63:043007, astro-ph/0008167.

Tegmark's Website (2003). `http://www.hep.upenn.edu/~max/`.

Theuns, T. et al. (2002). Detection of He II reionization in the SDSS quasar sample. *Astrophys. J.*, 574:L111–L114, astro-ph/0206319.

Tonry, J. L. et al. (2003). Cosmological results from high-z supernovae. *Astrophys. J.*, 594:1–24, astro-ph/0305008.

Trotta, R. (2001). Cosmic microwave background anisotropies: Dependence on initial conditions. Diploma Thesis of the Federal Institute of Technology (ETHZ), Zurich, Switzerland, available from the website: `http://theory.physics.unige.ch/~trotta`.

Trotta, R. & Durrer, R. (2004). Testing the paradigm of adiabaticity. In Ruffini et al., editors, *Proceedings of the X Marcel Grossman Meeting, 20-26 July 2003, Rio de Janeiro*. astro-ph/0402032. To appear.

Trotta, R. & Hansen, S. H. (2004). Observing the helium abundance with CMB. *Phys. Rev.*, D69:023509, astro-ph/0306588.

Trotta, R., Riazuelo, A., & Durrer, R. (2001). Reproducing cosmic microwave background anisotropies with mixed isocurvature perturbations. *Phys. Rev. Lett.*, 87:231301, astro-ph/0104017.

Trotta, R., Riazuelo, A., & Durrer, R. (2003). The cosmological constant and general isocurvature initial conditions. *Phys. Rev.*, D67:063520, astro-ph/0211600.

Uzan, J.-P. (1998). Dynamics of relativistic interacting gases: from a kinetic to a fluid description. *Class. Quant. Grav.*, 15:1063–1088, gr-qc/9801108.

Uzan, J.-P. (2003). The fundamental constants and their variation: Observational status and theoretical motivations. *Rev. Mod. Phys.*, 75:403, hep-ph/0205340.

Valiviita, J. & Muhonen, V. (2003). Correlated adiabatic and isocurvature CMB fluctuations in the wake of WMAP. *Phys. Rev. Lett.*, 91:131302, astro-ph/0304175.

Verde, L. et al. (2002). The 2dF galaxy redshift survey: The bias of galaxies and the density of the Universe. *Mon. Not. Roy. Astron. Soc.*, 335:432, astro-ph/0112161.

Verde, L. et al. (2003). First year Wilkinson Microwave Anisotropy Probe (WMAP) observations: Parameter estimation methodology. *Astrophys. J. Suppl.*, 148:195, astro-ph/0302218.

Vilenkin, N. Y. & Smorodinskii, Y. A. (1964). *Sov. Phys. JETP*, 19:1209.

Wandelt, B. D., Hivon, E., & Gorski, K. M. (2001). Cosmic microwave background anisotropy power spectrum statistics for high precision cosmology. *Phys. Rev.*, D64:083003, astro-ph/9808292.

Wang, X.-m., Tegmark, M., & Zaldarriaga, M. (2002). Is cosmology consistent? *Phys. Rev.*, D65:123001, astro-ph/0105091.

Webb, J. K. et al. (2001). Further evidence for cosmological evolution of the fine structure constant. *Phys. Rev. Lett.*, 87:091301, astro-ph/0012539.

Webb, J. K., Murphy, M. T., Flambaum, V. V., & Curran, S. J. (2003). Does the fine structure constant vary? A third quasar absorption sample consistent with varying alpha. *Astrophys. Space Sci.*, 283:565, astro-ph/0210531.

Weller, J. & Albrecht, A. (2002). Future supernovae observations as a probe of dark energy. *Phys. Rev.*, D65:103512, astro-ph/0106079.

Wetterich, C. (1988). Cosmology and the fate of dilatation symmetry. *Nucl. Phys.*, B302:668.

White, M. J., Gelmini, G., & Silk, J. (1995). Structure formation with decaying neutrinos. *Phys. Rev.*, D51:2669–2676, astro-ph/9411098.

Will, C. M. (2001). The confrontation between general relativity and experiment. *Living Rev. Rel.*, 4:4, gr-qc/0103036.

Wittman, D. M., Tyson, J. A., Kirkman, D., Dell'Antonio, I., & Bernstein, G. (2000). Detection of weak gravitational lensing distortions of distant galaxies by cosmic dark matter at large scales. *Nature*, 405:143–149, astro-ph/0003014.

Wong, Y. Y. Y. (2002). Analytical treatment of neutrino asymmetry equilibration from flavour oscillations in the early universe. *Phys. Rev.*, D66:025015, hep-ph/0203180.

Wright, E. L. et al. (1992). Interpretation of the CMB anisotropy detected by the COBE DMR. *Astrophys. J.*, 396:L13–L18.

Zaldarriaga, M. (1997). Polarization of the microwave background in reionized models. *Phys. Rev.*, D55:1822–1829, astro-ph/9608050.

Zaldarriaga, M. & Harari, D. D. (1995). Analytic approach to the polarization of the cosmic microwave background in flat and open universes. *Phys. Rev.*, D52:3276–3287, astro-ph/9504085.

Zaldarriaga, M. & Seljak, U. (1997). An all-sky analysis of polarization in the microwave background. *Phys. Rev.*, D55:1830–1840, astro-ph/9609170.

Zaldarriaga, M. & Seljak, U. (2000). CMBFAST for spatially closed universes. *Astrophys. J. Suppl.*, 129:431–434, astro-ph/9911219.

Zaldarriaga, M., Seljak, U., & Bertschinger, E. (1998). Integral solution for the microwave background anisotropies in non-flat universes. *Astrophys. J.*, 494:491, astro-ph/9704265.

Zaldarriaga, M., Spergel, D. N., & Seljak, U. (1997). Microwave background constraints on cosmological parameters. *Astrophys. J.*, 488:1–13, astro-ph/9702157.

Zel'dovich, Y. B. (1972). A hypothesis, unifying the structure and entropy of the universe. *Mon. Not. Roy. Astron. Soc.*, 160:1–3.

Zentner, A. R. & Walker, T. P. (2002). Constraints on the cosmological relativistic energy density. *Phys. Rev.*, D65:063506, astro-ph/0110533.