BOSONIZATION: HOW TO MAKE IT WORK FOR YOU IN CONDENSED MATTER*,**

R. SHANKAR

Center for Theoretical Physics, P.O. Box 6666
Yale University
New Haven, CT 06511 U.S.A.

An elementary introduction to Abelian bosonization is provided here. It is shown that although it is applicable to relativistic Dirac fermions, there are many examples of problems in condensed matter theory which are described (with respect to some characteristics) by just such objects. Examples considered here are the uniform and random Ising models, and the Hubbard model.

PACS numbers: 71.10.Hf, 11.10.Kk

1. Introduction

I thank the organizers for inviting me to this wonderful haven. Just as the tour book said, we came to Poland as visitors and will leave as friends.

Some of the work described here was due to Claude Itzykson of Saclay, who passed away recently. It was my good fortune to have known this warm and charismatic gentleman. I fondly dedicate these lectures to Claude's memory.

Now for our topic proper. Bosonization refers to the possibility of describing a theory of relativistic Dirac fermions obeying standard anticommutation rules by a boson field theory. While this may be possible in all dimensions, it has so far proved most useful only in $d = 1$ for the simple reason that only in this case is the bosonic version of the given fermionic theory local and simple, and often simpler that the Fermi theory.

---

* Presented at the XXXV Cracow School of Theoretical Physics, Zakopane, Poland June 4–14, 1995.

** Lectures given at the BCSPIN School, Katmandu, May 1991. Reprinted after minor changes with the kind permission of World Scientific from *Current Topics in Condensed Matter and Particle Physics*, Editors J.Pati, Q.Shafi and Yu Lu, World Scientific (1993).

To my knowledge bosonization, as described here, was first carried out by Lieb and Mattis [1] in their exact solution of the Luttinger model. Later Luther and Peschel [2] showed how to use it to find asymptotic (low momentum and energy) correlation function for more generic interacting Fermi systems. It was independently discovered in particle physics by Coleman [3] and further developed by Mandelstam [4]. Much of what I know and use is inspired by the work of Luther and Peschel.

Before getting into any details, I would first like to answer two questions and thereby provide you with an overview. First, if bosonization applies only to relativistic Dirac fermions, why is it of any interest to condensed matter theory where relativity is never considered? Second, what is the magic by which bosonization helps us tame interacting field theories?

As for the first question, there are two ways in which Dirac fermions enter condensed matter physics. The first is in the study of two-dimensional Ising models. If we use the transfer matrix approach and convert the classical problem on an N-by-N lattice to a quantum problem in one dimension we end up with a 2^N dimensional Hilbert space, with a Pauli matrix at each of N sites. The two dimensions at each site represent the two-fold choice of values open to the Ising spins. Consider now a spinless fermion degree of freedom at each site. Here too we have two choices: the fermion state is occupied or empty. There is some need for cleverness in going from the Pauli matrix problem to the fermion problem since Pauli matrices commute at different sites while fermions anticommute; this was provided by Jordan and Wigner. At the critical point the fermions had to be massless and relativistic: massless since the system is critical and relativistic since one regains all the symmetries of the continuum at the critical point. All this will be made precise shortly.

The second way in which Dirac fermions arise is in the study of spinless fermions on a linear lattice. Consider the noninteracting case with just hopping:

$$H = -\sum_{n=-\infty}^{\infty} \psi^\dagger(n)\psi(n+1) + \text{h.c.}. \tag{1}$$

In the above, the spinless fermion field obeys the standard anticommutation rules

$$\{\psi^\dagger(n), \psi(m)\} = \delta_{mn} \tag{2}$$

with all other anticommutators vanishing.

We switch to plane waves (with system volume equal to unity)

$$\psi(n) = \int_{-\pi}^{\pi} \frac{dk}{2\pi} \psi(k) e^{ikn} \tag{3}$$

obeying

$$\{\psi^\dagger(k), \psi(q)\} = 2\pi\delta(k-q)\,. \tag{4}$$

In inverting relation (3) you must use $2\pi\delta(0) = N$, the volume of the system. The Hamiltonian separates nicely in terms of these operators:

$$H = -\int_{-\pi}^{\pi} \frac{dk}{2\pi} \cos k \psi^\dagger(k)\psi(k)\,. \tag{5}$$

In the ground state we must fill all negative energy modes, that is, states between $\pm k_F$, where $k_F = \pi/2$. If we now need to study the low energy properties of the system, we can focus on the modes near just the Fermi points. We find that they have $E = \pm k$, where k is measured from the respective Fermi points. These are the two components of the massless Dirac field. Any interaction between the primordial fermions can be written in terms of these two components.

Next, we ask how bosonization can make life easier. Say we have a problem where $H = H_0 + V$, where H_0 is the free Dirac Hamiltonian and V is a perturbation. Assume we can express all quantities of interest in terms of a power series in V. In the interaction picture the series will involve the correlation function of various operators evolving under H_0. Bosonization now tells us that the same series is reproduced by starting with $H = H_0^B + V^B$ where H_0^B is a massless free boson Hamiltonian and V^B is a bosonic operator which depends on V and is specified by the bosonization dictionary. Consider a special case where $V = \rho^2$ where $\rho = \psi^\dagger(x)\psi(x)$, the Dirac charge density. This is a quartic interaction in the Fermi language and obviously non trivial. But according to the dictionary, we must replace ρ by the bosonic operator $(1/\sqrt{\pi})\partial_x\phi$, ϕ being the boson field. Thus V is replaced by the *quadratic* interaction $(1/\pi)(\partial_x\phi)^2$. Thus the bosonic version is trivial! I must add that this is not always the case, a simple mass term in the Fermi language becomes the formidable interaction $\cos\sqrt{4\pi}\phi$.

Let us now begin our course. I will first remind you of some basic facts about massless fermions and bosons in one dimension. This will be followed by the bosonization dictionary that relates interacting theories in one language to the other. This will be followed by the first application: deriving the critical correlation function of the Ising model using bosonization. Next we will do the same for the random bond Ising model. The last section will be devoted to quantum Fermi systems linearized near their Fermi points as described earlier. All topics will be discussed essentially from first principles.

2. The massless Dirac fermion

In one dimension the Dirac equation

$$i\frac{\partial\psi}{\partial t} = H\psi\,, \tag{6}$$

will have

$$H = \alpha P + \beta m\,, \tag{7}$$

where P is the momentum operator, and α and β are Pauli matrices σ_3 and σ_2 respectively. Let us focus on the massless case. There is nothing to diagonalize now: $\psi_\pm$, the upper and lower components of ψ, called right and left movers, are decoupled. In terms of the field operators obeying:

$$\{\psi_\pm^\dagger(x)\,,\psi_\pm(y)\} = \delta(x-y)\,, \tag{8}$$

the second quantized Hamiltonian

$$\begin{aligned} H &= \int \psi^\dagger(x)(\alpha P)\psi(x)dx \\ &= \int \psi_+^\dagger(x)(-i\partial_x)\psi_+(x)dx + \int \psi_-^\dagger(x)(i\partial_x)\psi_-(x)dx\,. \end{aligned} \tag{9}$$

Going over to Fourier transforms

$$\psi_\pm(p) = \int_{-\infty}^{\infty} \psi_\pm(x)\mathrm{e}^{-ipx}dx\,, \tag{10}$$

obeying

$$\{\psi_\pm^\dagger(p)\,,\psi_\pm(q)\} = 2\pi\delta(p-q)\,, \tag{11}$$

we find

$$H = \int \psi_+^\dagger(p)\; p\; \psi_+(p)\frac{dp}{2\pi} + \int \psi_-^\dagger(p)\;(-p)\psi_-(p)\frac{dp}{2\pi}\,. \tag{12}$$

From the above, it is clear that the right/left movers have energies $E = \pm p$, respectively. The Dirac sea is thus filled with right movers of negative momentum and left movers with positive momentum.

Since the fields have trivial time evolution in this free field theory, we can write down the Heisenberg operators at all times:

$$\psi_{\pm}(x,\ t) = \int_{-\infty}^{\infty} \frac{dp}{2\pi} \psi_{\pm}(p) \mathrm{e}^{ip(x\mp t)}\,. \tag{13}$$

Notice the $\psi_{\pm}$ is a function of $x \mp t$ only.

Consider now the equal time correlation function in the ground state:

$$\langle \psi_{+}(x)\psi_{+}^{\dagger}(0)\rangle = \int_{-\infty}^{\infty} \frac{dp}{2\pi} \int_{-\infty}^{\infty} \frac{dq}{2\pi} \mathrm{e}^{ipx} \underbrace{\langle \psi_{+}(p)\, \psi_{+}^{\dagger}(q)\rangle}_{2\pi\delta(p-q)\theta(q)} = \int_{0}^{\infty} \frac{dp}{2\pi} \mathrm{e}^{ipx}\,. \tag{14}$$

We have used the fact that a right mover can be created only for positive momenta since the Dirac sea is filled with negative momentum particles. The last integral is ill defined at large momenta. So we introduce a convergence factor $\mathrm{e}^{-\alpha|p|}$ that cuts off large momenta. Unless the physics in question involves arbitrarily large momenta (and hence energies) this should not make any difference to anything: in other words, if at the end of the calculation we have a formula for some physical quantity, we will usually be able to set $\alpha = 0$. So now we have

$$\langle \psi_{+}(x)\psi_{+}^{\dagger}(0)\rangle = \int_{0}^{\infty} \frac{dp}{2\pi} \mathrm{e}^{-\alpha p} \mathrm{e}^{ipx} = \frac{1}{2\pi} \frac{1}{\alpha - ix}\,. \tag{15}$$

If we want the correlation function for unequal times, we just replace x by $x - t$ since we know that the right movers are functions of just this combination.

In the same way, we can show that

$$\langle \psi_{\pm}(x)\, \psi_{\pm}^{\dagger}(0)\rangle = \frac{\pm i/2\pi}{x \pm i\alpha}\,, \tag{16}$$

$$\langle \psi_{\pm}^{\dagger}(0)\, \psi_{\pm}(x)\rangle = \frac{\mp i/2\pi}{x \mp i\alpha}\,. \tag{17}$$

Note that

$$\langle \psi_{\pm}(x)\, \psi_{\pm}^{\dagger}(0) + \psi_{\pm}^{\dagger}(0)\, \psi_{\pm}(x)\rangle = \frac{\alpha/\pi}{x^2 + \alpha^2} \simeq \delta(x)\,, \tag{18}$$

where in the last equation we are considering the limit of vanishing α.

Besides the Fermi field, there are bilinears in the field that occur often. Let us look at some key ones. The current density j_{μ} has components

$$j_0 = \psi^{\dagger}\psi = \psi_{+}^{\dagger}(x)\psi_{+}(x) + \psi_{-}^{\dagger}(x)\psi_{-}(x) \tag{19}$$

$$j_1 = \psi^{\dagger}\alpha\psi = \psi_{+}^{\dagger}(x)\psi_{+}(x) - \psi_{-}^{\dagger}(x)\psi_{-}(x) \tag{20}$$

The axial current is given by $j^5_\mu = \varepsilon_{\mu\nu} j_\mu$. The last bilinear is the "mass term"

$$\bar{\psi}\psi = \psi^\dagger(x)\beta\psi(x) = -i\psi^\dagger_+(x)\psi_-(x) + i\psi^\dagger_-(x)\psi_+(x). \tag{21}$$

For later use, let us note that

$$\langle \bar{\psi}\psi(x)\, \bar{\psi}\psi(0)\rangle = \frac{1}{2\pi^2}\frac{1}{x^2+\alpha^2}. \tag{22}$$

The derivation of this result is left as an exercise. All you need are the anticommutation rules and the correlation functions from Eqs (16), (17).

We close the section on fermions with a digression on what are called Majorana fermions. These may be viewed as hermitian or real fermions. Let us call our Dirac field ψ_D and express in terms of two *Hermitian* fields ψ and χ:

$$\psi_D = \frac{\psi + i\chi}{\sqrt{2}}, \tag{23}$$

$$\psi^\dagger_D = \frac{\psi - i\chi}{\sqrt{2}}. \tag{24}$$

It is readily verified that

$$\{\psi(x), \psi(y)\} = \delta(x-y) \tag{25}$$

with a similar rule for χ and all other anticommutators vanishing. If we write the massive Dirac Hamiltonian in terms of the Majorana fields, defined above, we will get

$$\begin{aligned} H_D &= \int [\psi^\dagger_D(\alpha P + \beta m)\psi_D]dx \\ &= \tfrac{1}{2}\int [\psi^T(\alpha P + \beta m)\psi] + [\chi^T(\alpha P + \beta m)\chi] \\ &\quad - i[\psi^T(\alpha P + \beta m)\chi] + i[\chi^T(\alpha P + \beta m)\psi]. \end{aligned} \tag{26}$$

You may check that the cross terms involving β in fact add to zero. To fully decouple the Majorana fields, we change the representation of the α matrix so it equals Pauli's σ_1. It is now easy to verify (upon doing an integration by parts) that the cross terms do indeed drop out and the Hamiltonian is a sum of two Majorana Hamiltonians. Thus one free Dirac fermion equals two Majorana fermions just as one charged scalar field equals two real fields (not just in degrees of freedom, but at the level of H).

3. Free massless scalar field

The Hamiltonian for a massless scalar field is

$$H_B = \tfrac{1}{2}\int (\Pi^2 + (\nabla\phi)^2)dx\,, \tag{27}$$

where Π and ϕ obey

$$[\phi(x), \Pi(y)] = i\delta(x-y)\,. \tag{28}$$

The Schrödinger operators are expanded as follows:

$$\phi(x) = \int_{-\infty}^{\infty} \frac{dp}{2\pi\sqrt{2|p|}} \left[\phi(p)e^{ipx} + \phi^\dagger(p)e^{-ipx}\right] e^{-\alpha|p|}\,, \tag{29}$$

$$\Pi(x) = \int_{-\infty}^{\infty} \frac{dp|p|}{2\pi\sqrt{2|p|}} \left[-i\phi(p)e^{ipx} + i\phi^\dagger(p)e^{-ipx}\right] e^{-\alpha|p|}\,. \tag{30}$$

Due to the convergence factors, ϕ and Π will obey

$$[\phi(x), \Pi(y)] = \frac{i\alpha/\pi}{\alpha^2 + x^2} \simeq i\delta(x-y)\,. \tag{31}$$

The Hamiltonian now takes the form:

$$H = \int_{-\infty}^{\infty} \frac{dp}{2\pi}|p|\phi^\dagger(p)\phi(p)\,, \tag{32}$$

as you may show for yourself. We now introduce right and left movers as per

$$\begin{aligned}
\phi_\pm(x) &= \tfrac{1}{2}\left[\phi(x) \mp \int_{-\infty}^{x} \Pi(x')dx'\right] \\
&= \tfrac{1}{2}\int_{-\infty}^{\infty} \frac{dp}{2\pi\sqrt{2|p|}} e^{-\alpha|p|}\left[\phi(p)\left(1 \mp \frac{|p|}{p}\right)e^{ipx} + \text{h.c.}\right] \\
&= \pm\int_{0}^{\pm\infty} \frac{dp}{2\pi\sqrt{2|p|}} e^{ipx}e^{-\alpha|p|}\,.
\end{aligned} \tag{33}$$

I leave it to you to verify that

$$[\phi_\pm(x), \phi_\pm(y)] = \tfrac{i}{4}\varepsilon(x-y)\,, \tag{34}$$

$$[\phi_+(x), \phi_-(y)] = \tfrac{i}{4}\,. \tag{35}$$

(The studious reader will note that due to the convergence factors, a rounded out step function will arise in place of $\varepsilon(x-y)$, and this will become a step function as $\alpha \to 0$.) Note also that if we use the Heisenberg equations of motion for $\phi(p)$ and $\phi^\dagger(p)$ we will find that $\phi_\pm$ are functions only of $x \mp t$.

We must next work out some correlation functions in this theory. It is claimed that

$$G_\pm(x) = \langle \phi_\pm(x)\phi_\pm(0) - \phi_\pm^2(0)\rangle \tag{36.a}$$

$$= \tfrac{1}{4\pi}\ln\frac{\alpha}{\alpha \mp ix} \tag{36.b}$$

$$G(x) = \langle \phi(x)\phi(0) - \phi^2(0)\rangle \tag{37.a}$$

$$= \tfrac{1}{4\pi}\ln\frac{\alpha}{\alpha \mp ix}\,. \tag{37.b}$$

I will now establish one of them leaving the rest as exercises. Consider

$$G_+ = \int_0^\infty \frac{dp}{\sqrt{4\pi|p|}} \int_0^\infty \frac{dq}{\sqrt{4\pi|q|}} \langle(\phi(p)\phi^\dagger(q)\rangle(e^{ipx}-1)e^{-\alpha p}$$

$$= \int_0^\infty \frac{dp}{4\pi|p|}(e^{ipx}-1)e^{-\alpha p}$$

$$= \tfrac{1}{4\pi}\ln\frac{\alpha}{\alpha - ix}\,, \tag{38}$$

where the last line comes from looking up a table of integrals. If you cannot find this particular form of the result, I suggest you first differentiate both sides with respect to x, thereby eliminating the $1/|p|$ factor. Now the integral is easily shown to be $i/(4\pi(\alpha - ix))$. Next integrate this result with respect to x, using as a boundary condition $G_+(0) = 0$ to get the quoted result.

Finally we consider a class of operators one sees a lot of in two (space-time) dimensional theories. These are exponentials of the scalar field. Consider first

$$G_\beta \equiv \langle e^{i\beta\phi(x)}\, e^{-i\beta\phi(0)}\rangle\,. \tag{39}$$

First note that for the correlator to be nonzero, the sum of the factor multiplying ϕ in the exponentials has to vanish. This is so because the theory

(the Hamiltonian of the massless scalar field) is symmetric under a constant shift in ϕ. To evaluate this correlator, we need the following identity:

$$e^A + e^B =: e^{A+B} : \exp\left(\left\langle AB + \frac{A^2 + B^2}{2}\right\rangle\right), \tag{40}$$

where the normal ordered operator $: A :$ has all its destruction operators to the right and creation operators to the left. In the case of $: e^{\phi} :$, the vacuum expectation value is unity since in the exponential series, only the unity fails to annihilate the vacuum acting to the right or left. If you want to amuse yourself by proving Eq. (40) here is a possible route. Start with the more familiar identity (which we will not prove):

$$\begin{aligned} e^{A+B} &= e^A e^B e^{-(1/2)[A,B]} \\ &= e^B e^A e^{(1/2)[A,B]} \end{aligned} \tag{41}$$

provided $[A, B]$ commutes with A and B. Using this, first write $e^A = e^{A^+ + A^-}$, where $A^{\pm}$ are the creation and destruction parts of A, in normal ordered form. Now turn to e^{A+B}, separate the exponentials using the identity above, next normal order each part using this formula again and finally normal order the whole thing. (The last step is needed because $:A::B:$ is not itself normal ordered.) We now use Eq. (40) to evaluate G_β:

$$\begin{aligned} G_\beta &= \langle : \exp\left(i\beta(\phi(x) - \phi(0))\right) : \rangle \\ &\quad \times \exp\left(\beta^2\left[\left\langle \phi(x)\phi(0) - \frac{\phi^2(0) + \phi^2(x)}{2}\right\rangle\right]\right) \\ &= \exp\left(\beta^2 \frac{1}{4\pi} \ln \frac{\alpha^2}{\alpha^2 + x^2}\right) \\ &= \left(\frac{\alpha^2}{\alpha^2 + x^2}\right)^{\beta^2/4\pi} \end{aligned} \tag{42}$$

Notice two things. First, by varying β we can get operators with a continuum of power law decays in correlation. Next, as we send α to 0, the correlator vanishes. To avoid this we must begin with operators suitably boosted or renormalized. The thing to do in the above example is to consider:

$$[e^{i\beta\phi}]_R = \alpha^{-\beta^2/4\pi} e^{i\beta\phi}. \tag{43}$$

This operator will have finite correlations in the limit of zero α: if we give it less of a boost it dies, more and it blows up.

One can similarly show that

$$\langle e^{i\beta\phi_\pm(x)} e^{-i\beta\phi_\pm(0)} \rangle = \left(\frac{\alpha}{\alpha \mp ix}\right)^{\beta^2/4\pi}. \tag{44}$$

4. Bosonization dictionary

So far we have dealt with massless Fermi and Bose theories and the behaviour of various correlation functions in each. Now we are ready to discuss the rules for trading the Fermi theory for the Bose theory. The most important formula is this:

$$\psi_{\pm}(x) = \frac{1}{\sqrt{2\pi\alpha}} e^{\pm i\sqrt{4\pi}\phi_{\pm}(x)}. \tag{45}$$

This means that any correlation function of the Fermi field, calculated in the Fermi vacuum with the given (α) cut-off, is reproduced by the correlator of the bosonic operator given in the right hand side, if computed in the bosonic vacuum with the same momentum cut-off. Given this operator equivalence, we can replace any interaction term made out of the fermionic field by the corresponding bosonic counterpart. Sometimes, this will require some care, but this is the general idea.

There are several ways to convince you of the correctness of the above master formula. First consider the correlation

$$\langle \psi_+(x)\psi_+^{\dagger}(0)\rangle = \frac{1}{2\pi}\frac{1}{\alpha - ix}. \tag{46}$$

Let us see this reproduced by the bosonic version:

$$\begin{aligned}
&\Big\langle \frac{1}{\sqrt{2\pi\alpha}} e^{i\sqrt{4\pi}\phi_+(x)} \frac{1}{\sqrt{2\pi\alpha}} e^{-i\sqrt{4\pi}\phi_+(0)} \Big\rangle \\
&= \frac{1}{2\pi\alpha} \langle : e^{i\sqrt{4\pi}\phi_+(x)} e^{-i\sqrt{4\pi}\phi_+(0)} : \rangle e^{4\pi\langle \phi_+(x)\phi_+(0) - \phi_+^2\rangle} \\
&= \frac{1}{2\pi\alpha} e^{4\pi G_+(x)} \\
&= \frac{1}{2\pi\alpha}\frac{\alpha}{\alpha - ix}
\end{aligned} \tag{47}$$

In the above we have used the normal ordering formula Eq. (40), the definition of G_+ from Eq. (36.a) and its actual value from Eq. (36.b).

It is possible to verify at an operator level that the bosonized version of the Fermi field obeys all the anticommutation rules (with delta functions of width α). I leave this to the more adventurous ones among you. Instead I will now consider some composite operators and show the care needed in dealing with their bosonization. The first of these is

$$\overline{\psi}\psi = -\frac{1}{\pi\alpha}\cos\sqrt{4\pi}\phi. \tag{48}$$

The proof involves just the use of Eq. (41) and goes as follows:

$$\begin{aligned}\overline{\psi}\psi &= -i\psi_+^\dagger(x)\psi_-(x) + \text{h.c.} \\ &= \frac{1}{2\pi\alpha}\left[e^{-i\sqrt{4\pi}\phi_+(x)}e^{-i\sqrt{4\pi}\phi_-(x)}(-i) + \text{h.c.}\right] \\ &= \frac{1}{2\pi\alpha}\left(e^{-i\sqrt{4\pi}\phi(x)}e^{\frac{1}{2}4\pi(-1)\frac{i}{4}}(-i) + \text{h.c.}\right) \\ &= -\frac{1}{\pi\alpha}\cos\sqrt{4\pi}\phi\,. \end{aligned} \tag{49}$$

The factor $i/4$ in the exponent arises from the commutator of the right and left movers, Eq. (35).

In the above manipulations we brought together two operators at the same point. Each one has been judiciously scaled to give sensible matrix elements (neither zero nor infinite) acting on the vacuum. There is no guarantee that a product of two such well behaved operators at the same point is itself well behaved. A simple test is to see if the product has a finite matrix element in the vacuum as the points approach each other. In the above example, this was the case: in fact the mean value of the composite operator is zero since they create and destroy different (right or left moving) fermions. This is not the case for the next item: the operator $\psi_+^\dagger(x)\psi_+(x)$. We define it by a limiting process as follows:

$$\begin{aligned}\psi_+^\dagger(x)\psi_+(x) &= (\lim_{x\to 0})\frac{1}{2\pi\alpha}e^{-i\sqrt{4\pi}\phi_+(x)}e^{i\sqrt{4\pi}\phi_+(0)} \\ &= (\lim_{x\to 0})\frac{1}{2\pi\alpha} : e^{-i\sqrt{4\pi}\phi_+(x)}e^{i\sqrt{4\pi}\phi_+(0)} : e^{4\pi G_+(x)} \\ &= \frac{i}{2\pi(x+i\alpha)} : 1 - i\sqrt{4\pi}\frac{\partial\phi_+}{\partial x}x + \cdots : \frac{i\alpha}{x+i\alpha} \\ &= \frac{i}{2\pi x} + \frac{1}{\sqrt{\pi}}\frac{\partial\phi_+}{\partial x} + \cdots\,. \end{aligned} \tag{50}$$

The above manipulations need some explanation. First observe that we perform a Taylor expansion only within the normal ordering symbols. This is because only the normal ordered operators have nice (differentiable) matrix elements. Thus terms higher order in x and sitting within the symbol are indeed small and can be dropped as $x \to 0$. Consider next the $x + i\alpha$ in the denominator. Is it permissible to drop the α in comparison to x, even though x itself is being sent to 0? Yes, we must always treat any distance x in the continuum theory as being smaller than α, which is essentially a spatial short-distance cut-off. Finally note that the density operator in question has an infinite c-number part, we have displayed in front. This

reflects the fact that the vacuum density of right movers is infinite due to the Dirac sea. If we define a normal ordered density, *i.e.*, take away the singular vacuum average from it, we obtain

$$:\psi_+^\dagger(x)\psi_+(x):=\frac{1}{\sqrt{\pi}}\frac{\partial\phi_+}{\partial x}. \tag{51}$$

A similar result obtains for the left-mover density. Combining the two we get some very famous formulae in bosonization:

$$j_0=\frac{1}{\sqrt{\pi}}\frac{\partial\phi}{\partial x} \tag{52}$$

$$j_1=\frac{1}{\sqrt{\pi}}\frac{\partial(\phi_+-\phi_-)}{\partial x}$$

$$=-\frac{\Pi}{\sqrt{\pi}}. \tag{53}$$

We close this section with what is a very useful but odd looking relation:

$$\left[\frac{-1}{\pi\alpha}\cos\sqrt{4\pi}\phi\right]^2=-\frac{1}{\pi}(\frac{\partial\phi_+}{\partial x})^2+\frac{1}{2\pi^2\alpha^2}\cos\sqrt{16\pi}\phi+c\text{-numbers}. \tag{54}$$

To derive this result, you have to point-split the two cosines as we did in the density problem, write each as a sum of two exponentials. You must then combine exponentials with opposite exponents within the normal ordering symbol only, going this time to second order in the Taylor series since it is an even function we are expanding. The first term on the right hand side comes from doing this. The second comes from the naive combination of exponentials of the same sign, allowed since the product has no vacuum expectation value. In the field theory literature you will not see the second term mentioned. The reason is that at weak coupling this operator is highly irrelevant. By this I mean that if you find its correlation functions they will vanish despite the negative powers of α in front of it: though these boost it up, the boost is not enough. You can check this by considering its two-point function. The reason for keeping this is that in the presence of strong interactions the importance of this operator is decided by its correlation functions in the strongly interacting (as compared to the present noninteracting) theory. This is what happens in the XXZ chain.

5. Application I: uniform Ising model correlations at criticality

We now pass from this rather sterile business of deriving bosonization to actually using it. Of the countless applications, I have chosen a few that

I am most familiar with. While my treatment of the subject will not be exhaustive, it should prepare you to read any more material dealing with the subject.

Consider the Ising model in $d = 2$ with a partition function:

$$Z = \sum_{s=\pm 1} \exp\left(K \sum_{\langle ij \rangle} s_i s_j\right), \tag{55}$$

where $\langle i, j \rangle$ tells us that the Ising spins $s_i = \pm 1$ and $s_j = \pm 1$ are nearest neighbors on the square lattice. The sum in the exponent is a sum over bonds of the square lattice.

The correlation function

$$G(r) = \langle s_r s_0 \rangle, \tag{56}$$

where 0 is the origin and r a point a distance r away is defined as the following thermal average:

$$G(r) = \frac{\sum_{s=\pm 1} s_r s_0 \exp\left(K \sum_{\langle ij \rangle} s_i s_j\right)}{\sum_{s=\pm 1} \exp\left(K \sum_{\langle ij \rangle} s_i s_j\right)}. \tag{57}$$

It generally falls exponentially with r, except at the critical point where it falls like a power, which is known to be 1/4. This power is universal, *i.e.*, it will not change if we make some change in the interaction, such as adding a second neighbor term or making it anisotropic. But it will be unstable to adding a term odd in the spin variables, since this changes the symmetry or to the addition of a long range interaction which alters the effective dimensionality. (In the anisotropic case the distance r will naturally be given in terms of the coordinate differences and an anisotropic metric.) This exponent of 1/4 is rather difficult to derive and the reason will be more or less clear to you as we go along. I will now describe a trick due to Itzykson and Zuber [5] which uses bosonization to get around this.

Let us begin with a review of the transfer matrix formalism. For this purpose consider the model in $d = 1$. The lattice now is an array of N dots which we imagine to be vertical, to represent discrete units of time and to be connected so that the $N + 1$-th point lies next to the first, *i.e.*, with periodic boundary conditions. Consider

$$\begin{aligned} Z &= \sum_{s_i} \exp\left(\sum_i K(s_i s_{i+1} - 1)\right) \\ &= \sum_{s_i} \prod_i \exp\left(K(s_i s_{i+1} - 1)\right) \\ &= \sum_{s_i} T_{s_1 s_2} T_{s_2 s_3} \cdots T_{s_N s_1}, \end{aligned} \tag{58}$$

where we have subtracted a constant K from the bond energy, and, in the last step, introduced a 2×2 matrix T whose rows and columns are labeled by a pair of spins and whose element $T_{ss'}$ equals the Boltzmann weight associated with a pair of neighboring spins in the state s, s'. Thus $T_{++} = T_{--} = 1$, $T_{+-} = T_{-+} = \exp(-2K)$. In terms of this matrix

$$T = I + \mathrm{e}^{-2K}\sigma_1 , \tag{59}$$

$$Z = \mathrm{Tr}\, T^N . \tag{60}$$

Let us rewrite T as follows. Consider the identity

$$\begin{aligned} \mathrm{e}^{K^*\sigma_1} &= \cosh K^* + \sinh K^*\sigma_1 \\ &= \cosh K^*(I + \tanh K^*\sigma_1) . \end{aligned} \tag{61}$$

If we choose

$$\tanh K^* = \mathrm{e}^{-2K} \tag{62}$$

we see that up to a prefactor $\cosh K^*$,

$$T = \mathrm{e}^{K^*\sigma_1} . \tag{63}$$

We will drop this prefactor henceforth. This prefactor will modify the free energy by a term which is analytic except at infinite K^*. It will not affect the behaviour of correlation functions. Note that K^*, called the dual of K, is large when K is small and vice versa.

From Eq. (60) we see that finding Z reduced to finding the eigenvalues of T, which is readily done in this simple 2×2 case. Usually one has a more modest goal. One wants just f, the free energy per site in the infinite N limit. Then we need just the largest eigenvalue λ_0:

$$\begin{aligned} Z &= \mathrm{Tr}\, T^N \\ &= \sum_{i=1}^{2} \lambda_i^N \\ &\simeq \lambda_0^N \quad (\text{as } N \to \infty) \end{aligned} \tag{64}$$

$$\begin{aligned} f &= (\lim_{N\to\infty}) \ln \frac{Z}{N} \\ &= \ln \lambda_0 . \end{aligned} \tag{65}$$

For later reference let us note that in the present case, the eigenvalues are $\exp(\pm K^*)$ and the corresponding eigenvectors are

$$|0\rangle, |1\rangle = \tfrac{1}{\sqrt{2}}(1 , \pm 1) . \tag{66}$$

Consider next the correlation function $\langle s_i s_j \rangle$ for $i > j$. I claim that

$$\langle s_i s_j \rangle = \frac{\mathrm{Tr}\, T^{N-i} \sigma_3 T^{i-j} \sigma_3 T^j}{\mathrm{Tr}\, T^N} . \tag{67}$$

To see the correctness of this, look at the numerator. Retrace our derivation by introducing a complete set of σ_3 eigenstates between every factor of T. Reading from right to left, we get just the Boltzmann weights till we get to site i. There the σ_3 acting on its eigenstates, gives s_i, the value of the spin there. Then we proceed as usual to j, repeat this and go to the N-th site. This clearly agrees with Eq. (57). Let us rewrite this another way. Define Heisenberg operators

$$\sigma_a(j) = T^{-j} \sigma_a T^j , \qquad a = x, y, z . \tag{68}$$

In terms of these

$$\langle s_i s_j \rangle = \frac{\mathrm{Tr}\, T^N \sigma_3(i) \sigma_3(j)}{\mathrm{Tr}\, T^N} . \tag{69}$$

Consider now the limit as $N \to \infty$, i and j fixed. If we write T as

$$T = \sum_i |i\rangle \langle i| \lambda_i , \tag{70}$$

then for large N

$$T^N \simeq |0\rangle \langle 0| \lambda_0^N \tag{71}$$

and in the infinite volume limit

$$\langle s_i s_j \rangle = \langle 0 | \sigma_3(i) \sigma_3(j) | 0 \rangle . \tag{72}$$

(To verify this, combine Eqs. (69), (71) and take the trace in the eigenbasis of T.) For the case $j > i$, we will get the operators in the other order. In general then

$$\langle s_i s_j \rangle = \langle 0 | \tau(\sigma_3(i) \sigma_3(j)) | 0 \rangle , \tag{73}$$

where the time-ordering symbol τ will order the operators with time increasing from the right to left.

Now replace $\sigma_3(j)$ by the unit operator in the above derivation. It follows that the mean magnetization is

$$\langle s_i \rangle = \langle 0 | \sigma_3(0) | 0 \rangle . \tag{74}$$

In our example, $|0\rangle$ is the eigenket of σ_1 so that there is no mean magnetization. The only exception is at zero temperature or zero K^*: now the

eigenvalues are equal and we can form from the degenerate space a fully ordered state.

Although the preceding analysis and in particular Eq. (73) suggests that everything depends on just the ground state, a knowledge of all states is required even in the infinite volume limit. Going to Eq. (73) for the case $i > j$, let us insert the complete set of (two) eigenvectors of T between the Pauli matrices. When we insert $|0\rangle\langle 0|$ we get zero (since there is no magnetization) and when we insert $|1\rangle\langle 1|$ we get

$$\begin{aligned}\langle s_j s_i\rangle &= \langle 0|T^{-i}\sigma_3(0)T^{i-j}|1\rangle\langle 1|\sigma_3(0)T^j|0\rangle \\ &= \left(\frac{\lambda_1}{\lambda_0}\right)^{i-j} |\langle 0|\sigma_3|1\rangle|^2 \\ &= \mathrm{e}^{-2K^*(i-j)}|\langle 0|\sigma_3|1\rangle|^2 . \end{aligned} \tag{75}$$

Notice that the correlation falls exponentially with distance with a coefficient $2K^*$, which is just the gap to the first excited state of the Hamiltonian H defined by $T = \mathrm{e}^{-H}$ which in our example is $-K^*\sigma_1$.

This simple example has revealed most of the general features of the transfer matrix formalism in any context. The only difference is that for a bigger matrix, there will be a sum over decaying exponentially. Asymptotically the gap to the first excited state will rule.

Let us now move on to an $N \times M$ lattice $d = 2$. Now we have M copies of the column of N dots. We think of the horizontal dots as a spatial lattice and the vertical dots as discrete time just as in $d = 1$. The Boltzmann weights involve couplings between a spin and its spatial neighbor (a new feature) and it temporal neighbor (as in $d = 1$). We must once again cook up a T matrix, acting on the direct product space of dimension 2^M with matrix elements that equal the corresponding Boltzmann weights. I state the result:

$$T = \exp\left(K^* \sum_{i=1}^{M} \sigma_1(i)\right) \exp\left(K \sum_{i=1}^{M} \sigma_3(i)\sigma_3(i)\right). \tag{76}$$

Sandwich this between a bra and a ket labeled by the eigenvalues of the M σ_3 matrices, the ket and bra describing the spins at a given slice and the next respectively. Acting to the right the second factor in T gives the Boltzmann factors within a row (each σ_3 is replaced by its eigenvalue) times the ket. We now see the first factor sandwiched between the direct product of M Ising spins operators. Each is identical to what we had in $d = 1$ and gives the Boltzmann factor for the M vertical bonds. With these changes in T, all the formulae are of the same form as in $d = 1$. For example $Z = \mathrm{Tr}\, T^M$ and $f = \ln Z/MN$. In the limit of infinite M, T can be replaced by the

projection to the dominant state. To get the full thermodynamic limit we must next look at this state in the limit $N \to \infty$. All correlation functions are given by the expectation values of Heisenberg operators in this state.

Note that unlike in $d = 1$, where T is the exponential of a simple operator, it is a product of two exponentials. We could of course combine them by brute force, but this would involve a never ending set of commutators involving the exponents. So we resort to the following trick invented by Fradkin and Susskind [6]. If K and K^* could both be made small, we could simply add the exponents dropping the commutators as being of higher order. However, as seen earlier, when one is small the other is large. Let therefore consider an anisotropic lattice with couplings K_t and K_x in the time and space directions. Let us choose the former very large and the latter very small, in particular let us choose

$$K_x = \tau\,, \tag{77}$$

$$K_t^* = \lambda\tau \quad (\tau \to 0)\,. \tag{78}$$

With this choice (called the τ-continuum limit)

$$T = \mathrm{e}^{-\tau H}\,, \tag{79}$$

$$H = -\lambda \sum \sigma_1(m) - \sum \sigma_3(m)\sigma_3(m+1)\,. \tag{80}$$

The idea of Fradkin and Susskind is that anisotropy will change the metric but not the exponent for decay or any other universal quantity.

Although the above H is translationally invariant and bilinear (at most) in its variables, we cannot solve this by fourier transformation. The reason is that σ's at different sites commute (resembling bosons) while those at the same site anticommute (resembling fermions). Thus their transforms will be neither bosonic nor fermionic operators. We follow here the magic of Lieb, Schulz and Mattis and trade the Pauli matrices for full fledged Fermi operators [7].

Let us define

$$\psi_1(n) = \frac{1}{\sqrt{2}} \left(\prod_{-\infty}^{n-1} \sigma_1 \right) \sigma_2(n)\,, \tag{81}$$

$$\psi_2(n) = \frac{1}{\sqrt{2}} \left(\prod_{-\infty}^{n-1} \sigma_1 \right) \sigma_3(n)\,. \tag{82}$$

We are now treating an infinite spatial lattice and the "string" of σ_1's comes from the far left to the point $n-1$. You should verify that the new fields are Majorana fermions obeying

$$\{\psi_i(n), \psi_j(m)\} = \delta_{ij}\delta_{mn}\,. \tag{83}$$

The role of the string is to introduce global anticommutation. This is all very good, but what will the string do to the Hamiltonian? A local (nearest neighbor) interaction can easily become nonlocal when we change to nonlocal variables. Luckily for us

$$\begin{aligned} H &= 2i\lambda \sum \psi_1(n)\psi_2(n) - 2i \sum \psi_1(n)\psi_2(n+1) \\ &= (1-\lambda) \sum (-i\psi_1(n)\psi_2(n) + i\psi_2(n)\psi_1(n)) \\ &+ \sum [-i(\psi_1(n)(\psi_2(n+1) - \psi_2(n)) + i(\psi_2(n)(\psi_1(n) - \psi_1(n-1))] \,. \end{aligned} \tag{84}$$

I urge you to verify these steps. I also urge you to invert the transformation expressing the Pauli matrices in terms of fermions. It is remarkable how the final theory is bilinear and local in fermions, and hence amenable to solution by Fourier transformation. Had there been a coupling to a magnetic field via a σ_3 term, you would find that the string does not cancel and that H contains a sum over sites of infinite strings of Fermi operators. This is why the Ising model in a field is yet unsolved. (In going to the last equation above, I have used the anticommuting nature of the fields to trade in the factor of 2 for two ways of writing the same operator.)

Let us imagine that our lattice has a spacing a. Define continuum operators $\psi_c = \psi/\sqrt{a}$ so that the lattice operators obey δ function anticommutation rules as $a \to 0$. In terms of these, we get in the continuum limit the following continuum Hamiltonian $H_c = H/a$:

$$H_c = \int \psi^T(\alpha P + \beta m)\psi dx \qquad m = \frac{1-\lambda}{a}, \tag{85}$$

where α is now σ_1, as we agreed earlier. Several points need explanation. First, the subscript c has been dropped on the fermion operators. Next, we have replaced differences by derivatives. This is allowed only if the operator has correlations long compared to the lattice size. Now the correlator has a range given in this free field theory by its inverse mass $a/(1-\lambda)$ since this is the lowest excitation over the vacuum. Thus we must be very close to $\lambda = 1$ before this continuum limit is valid. Let us assume that this is the case.

The quadratic Hamiltonian in Eq. (84) is easily diagonalized. By filling all the negative energy levels we get the ground state energy E_0. This energy (per unit spatial volume) is essentially the free energy per site of the square lattice model, a result I leave to you show. Let us turn instead to the correlation functions. Now it seems that in a free field theory this should be trivial. But it is not, the reason being that what we want is not the

two-point function of the Fermi field, but of the spin field, and the latter is a very complicated function of the former. Let us find the equal time correlation of two sites a distance n apart. (The power law for decay should be the same in all direction even though length scales are not.) Thus we need to look at

$$\begin{aligned}\langle 0|\sigma_3(0)\sigma_3(n)|0\rangle &= \langle 0|\sigma_3(0)\sigma_3(1)\sigma_3(1)\cdots\sigma_3(n-1)\sigma_3(n)|0\rangle \\ &= \langle 0|[2i\psi_1(0)\psi_2(1)\cdot 2i\psi_1(1)\psi_2(2) \\ &\cdots 2i\psi_1(n-1)\psi_2(n)]|0\rangle\,. \end{aligned} \tag{86}$$

Thus we find that the two-point function of spins is a $2n$ point function of fermions. This becomes very hard if we want the limit of large n: we must evaluate a determinant of arbitrarily large size. We are however presently interested in obtaining just the power law of the asymptotic decay of the spin–spin correlation. To this end we can follow the trick of Itzykson and Zuber. First note that apart from the end factors, $\psi_1(0)$ at the left and $\psi_2(n)$ at the right, we have the product over sites of

$$2i\psi_2(i)\psi_1(i) = \exp\left(\frac{i\pi}{2}[2i\psi_2(i)\psi_1(i)]\right)\,. \tag{87}$$

This equation follows from the fact that $2i\psi_2(i)\psi_1(i)$ is just like a Pauli matrix (with square unity). The exponent is just $\frac{i\pi}{2}\overline{\psi}\psi$. When we from the product over sites, it becomes a sum and in the continuum limit the integral of $\overline{\psi}\psi$ between 0 and R, where $R = na$ is the distance between the points in laboratory units. There is no simple way to evaluate

$$G(R) \simeq \langle 0\Big|\exp\left(\frac{i\pi}{2}\int_0^R \overline{\psi}(x)\psi(x)dx\right)\Big|0\rangle\,. \tag{88}$$

Consider now an auxiliary problem, where we have made two noninteracting copies of the Ising system, with spins called s and t, and associated Pauli matrices σ and τ and Majorana fermions ψ and χ. It is clear that

$$\begin{aligned}\langle s_n t_n s_0 t_0\rangle &= \langle s_n s_0\rangle\langle t_n t_0\rangle \\ &= G(n)^2 \end{aligned} \tag{89}$$

since the thermal averages proceed independently and identically for the two sectors. The trick is to find G^2 and then take the square root. Let us see how this works. First, we will be dealing with products of the following

terms:

$$\begin{aligned} 2i\psi_2\psi_1 2i\chi_2\chi_1 &= [-2i\chi_1\psi_1]\cdot[2i\chi_2\psi_2] \\ &= \exp\left(\frac{i\pi}{2}[2i\psi_1\chi_1 + 2i\psi_2\chi_2]\right) \\ &= \exp\left(i\pi : \psi_D^\dagger\psi_D :\right) . \end{aligned} \tag{90}$$

The last step need some explanation. Let us form a Dirac fermion

$$\psi_D = \frac{\psi + i\chi}{\sqrt{2}} \tag{91}$$

and consider its charge density:

$$\begin{aligned} \psi^\dagger\psi &= \tfrac{1}{2}(\psi_1 - i\chi_1)(\psi_1 + i\chi_1) + (1 \to 2) \\ &= i\psi_1\chi_1 + i\psi_2\chi_2 + 1 \qquad (92) \\ : \psi^\dagger\psi : &= i\psi_1\chi_1 + i\psi_2\chi_2 , \qquad (93) \end{aligned}$$

where I have used the fact that the vacuum density of the Dirac fermions is 1 per site, half for the right movers half for the left movers. (Recall that in momentum space half the states are filled, this translates into half per site in real space.) This brings us to our punch line: in view of the above

$$G^2(R) = \left\langle 0\middle| \exp\left(i\pi \int_0^R : \psi^\dagger(x)\psi(x) : dx\right)\middle|0\right\rangle \tag{94.a}$$

$$= \left\langle 0\middle| \exp\left(\int_0^R i\sqrt{\pi}\partial_x\phi dx\right)\middle|0\right\rangle \tag{94.b}$$

$$= \left\langle 0\middle| \exp\left(i\sqrt{\pi}\phi(R)\right)\exp\left(-i\sqrt{\pi}\phi(0)\right)\middle|0\right\rangle \tag{94.c}$$

$$= \frac{1}{R^{1/2}} . \tag{94.d}$$

Thus the beauty of the bosonization is that a nonlocal Green function in the Fermi language has become a local 2-point function in the bosonic language. Several points of explanation are needed. First, we have used Eq. (52) in going from the first to the second equation in the above sequence. Next we have used the fact that at the critical point the Fermi theory has no mass. Thus the bosonic ground state in which the bosonic correlator is evaluated is the free field vacuum. Lastly we have used

Eq. (42) to evaluate the desired two-point function. Taking the square root we find the desired decay law $G(R) \simeq R^{-1/4}$.

I have been careless about the end points where the product does not follow the pattern. If this is taken into account, one finds that we must use $\sin\sqrt{\pi}\phi$ in place of $e^{i\sqrt{\pi}\phi}$. This does not however change the critical exponent. If one tries the Itzykson–Zuber trick away from criticality one finds that one has to find the correlation function of the same operator but in the theory with an interaction $\cos 4\sqrt{\pi}\phi$ which is the bosonized version of the harmless looking mass term in the free Dirac theory.

To conclude, the following were the highlights of our derivation of $G(R)$:

- The critical theory of the Ising model in the extreme anisotropic τ continuum limit is a massless Majorana theory.
- The two-point function of spins a distance R apart is given by the average of the exponential of the integral of a fermion bilinear from 0 to R.
- By considering the square of G, we made the integrand referred to above into the normal ordered Dirac charge density.
- By bosonizing the latter into the derivative of ϕ, we got rid of the integral in the exponent and were left with just a two-point function coming from the end points of the integration.
- By evaluating this in the free field theory we found that G^2 falls off like $R^{-1/2}$. We then took the square root of this answer.

6. Application II: the random bond Ising model

Consider an Ising model in which the bonds are not uniformly strong with strength we have called K, but where the bonds are chosen randomly from an ensemble. This can happen in real systems due to vacancies, lattice imperfections and so on. We should therefore imagine that each sample is different and translationally noninvariant. The study of the $d = 2$ Ising model with such a complication was pioneered by Dotsenko and Dotsenko [8] (refered to as DD hereafter) in a very influential paper. I will now describe their work as well as further contributions by others myself included. You will see bosonization at work once more.

First let us understand what we want to calculate in a random system. The behaviour of an individual system with bonds chosen in a sample specific way from the ensemble of possibilities is not interesting, unless by luck we are dealing with a property which is sample independent. (The free energy per site in the infinite volume limit is one such object.) In general what one wants are physical quantities, first calculated sample by sample **and then** averaged over samples. This is called a *quenched average* and is

a lot more difficult problem than the *annealed average* in which one treats the bond strength as another statistical variable in thermal equilibrium, just like the Ising spins themselves. Which one should one use? If the bonds are frozen into some given values over the period of the measurements, we must take them as a fixed external environment and do the quenched average. If they fluctuate ergodically over the period of measurement, we must do the annealed average. The DD problem deals with quenched averages. In this case one must work with the averaged free energy $\overline{f}$ obtained by averaging $\ln Z$ over all samples. The temperature derivative of $\overline{f}$ gives the average internal energy so on. (As mentioned above, it is known that in the infinite volume limit, each sample will give the same f. This is not true for all quantities.) Similarly one can take two spins a distance R apart find the correlator G. This will depend on the absolute values of the coordinates since there is no translational invariance. However the ensemble average $\overline{G}$ will depend only on R. Besides these mean values, one can calculate the fluctuations around these mean values. Given the distance R and a temperature, there is a unique number $G(R)$ in a pure system describing the correlation. In our case there is probability $P(G(R))$ that $G(R)$ will have this or that value. We will return to this point in the end.

We have seen that the Ising model is described a noninteracting Majorana field theory. We can take this Hamiltonian and write Z as a path integral over Grassmann numbers as follows:

$$Z_M(K) = \int [d\psi] \exp\left(\int \overline{\psi}(\partial\!\!\!/ + m)\psi d^2x\right) . \tag{95}$$

(You must consider the lectures by professors Rajaraman or Mani [9, 10] to see how this is done.) In the above m, the mass is determined by the λ or equivalently the temperature. It vanishes at the critical temperature. We are assuming we are close enough to criticality for this continuum theory to be valid. Suppose now that the bonds, instead of being uniform, vary from point to point on the two-dimensional lattice, never swaying too far from criticality. This means that $m = m(x)$ varies with the two-dimensional coordinate x. $Z_M = Z_M(m(x))$ is thus a functional of $m(x)$. Let us assume that the probability distribution for m is a Gaussian at each site:

$$P(m(x)) = \prod_x \exp\left(\frac{-(m(x) - m_0)^2}{2g^2}\right) . \tag{96}$$

Hereafter we will focus on the case of zero mean: $m_0 = 0$. Thus each bond fluctuates symmetrically around the critical value. To find $\overline{f}$ we must calculate

$$\overline{f} = \int P(m(x)) \ln Z_M(m(x)) dm(x) . \tag{97}$$

Since we are averaging $\ln Z$ and not Z, we see that the problem is not as easy as that of adding an extra thermal variable $m(x)$. We circumvent this using what is called the replica trick. We use

$$\ln Z = (\lim_{n\to 0})\frac{Z^n - 1}{n}. \tag{98}$$

In what follows we will drop the minus one in the numerator since it adds a constant to the answer and also drop the factor of inverse n since it multiplies the answer by a factor without changing any of the critical properties. In short in Eq. (97) we can replace $\ln Z$ by Z^n (and of course send n to zero at the end). But Z^n is just the partition function of n replicas of the original model. Thus

$$\begin{aligned}\overline{f} &= \int \left[\prod_1^n d\psi_i\right] \exp\left(\int \sum_1^n \overline{\psi}_i(\partial\!\!\!/ + m(x))\psi_i d^2x\right) \exp\left(\frac{-m^2(x)}{2g^2}\right) dm(x) \\ &= \int \left[\prod_1^n d\psi_i\right] \exp\left(\int \sum_1^n \left[\overline{\psi}_i(\partial\!\!\!/)\psi_i + \frac{g^2}{2}\left(\sum_1^n \overline{\psi}_i\psi_i\right)^2\right] d^2x\right). \end{aligned} \tag{99}$$

Thus the randomness is traded in for an interacting but translationally invariant theory, called the n-component Gross–Neveu model [21]. The above is a shortened derivation of the DD result. It is understood that all calculations be performed for general n and that in any analytic expression where n occurs, it is set to zero. The value of the DD work is that it shows in detail that this crazy procedure is indeed doing the ensemble average we want to do.

Now DD proceed to deduce two results:

- The specific heat will have a "ln ln" divergence instead of the "ln" divergence of the pure system. To derive one must also explore the case $m_0 \neq 0$.
- The average two point function $\overline{G(R)}$ falls essentially like R^0 as compared to the $R^{-1/4}$ in the pure system.

While the first result seemed reasonable, the second did not. The reason is as follows. It can be shown that the Gross-Neveu model is essentially free field theory at large distances, the interactions falling logarithmically. It is known in that in such asymptotically free theories correlations are usually free-field like up to logarithms. Thus we can accept the change in the specific heat from "ln" to "ln ln" but not the change of the decay exponent from 1/4 to 0. It was however difficult to see what had gone wrong in the rather difficult calculation of DD who dealt with the formidable task of evaluating

an average like Eq. (94.b) (which is difficult enough in free field theory) in an interacting theory.

I decided to approach the problem a different way. Recall how, in the pure case, by considering the square of the correlation, we could convert the problem, via bosonization, to the evaluation of a two-point function. Let us try the same trick here. Consider any one sample with some given set of bonds. On it imagine making two copies of the Ising system. Then, following the previous reasoning from the last section

$$G^2(0,R,m(x)) = \frac{1}{Z_D(m(x))} \times \int [d\psi_D][d\overline{\psi}_D] \exp\left[\int d^2x \overline{\psi_D}(\partial\!\!\!/ + m(x))\psi_D\right] \exp\left[i\pi \int_0^R : \psi_D^\dagger \psi_D : dx\right]. \tag{100}$$

In the above, G remembers that one spin was at the origin and the other at R (in both copies). In principle one must move this pair over the lattice maintaining this separation. However this will be obviated by the replica averaging to follow since the replicated problem will be translationally invariant. Next, due to the doubling, we have a Dirac fermion. Finally the normalizing partition function downstairs is itself a functional of $m(x)$ which makes it hard to average G^2. So we do the following. Let us multiply top and bottom by Z^{n-1} and set $n = 0$. This gets rid of the denominator and adds $n-1$ copies upstairs. We have then

$$G^2(0,R,m(x)) = \int \Big[\prod_1^n d\psi_i d\overline{\psi_i}\Big] \exp\Big(\int \sum_1^n \overline{\psi}_i(\partial\!\!\!/ + m(x))\psi_i d^2x\Big) \times \exp\Big(i\pi \int_0^R \psi_1^\dagger \psi_1 dx\Big), \tag{101}$$

where the subscript 1 labels the species we started with and all fermions are understood to be Dirac. If we now do the gaussian average over $m(x)$, we just complete the squares on the mass term and obtain

$$\overline{G^2(R)} = \int \Big[\prod_1^n d\psi_i \overline{\psi_i}\Big] \exp\Big(\int \sum_1^n [\overline{\psi}_i(\partial\!\!\!/)\psi_i + \frac{g^2}{2}\Big(\sum_1^n \overline{\psi}_i\psi_i\Big)^2] d^2x\Big) \times \exp\Big(i\pi \int_0^R \psi_1^\dagger \psi_1 dx\Big). \tag{102}$$

Let us now bosonize this theory to obtain:

$$\overline{G^2(R)} = \int \prod_i d\phi_i \exp\left(\int d^2x \sum_1^n -\frac{1}{2}(\nabla\phi_i)^2 + \frac{g^2}{2\pi^2\alpha^2}\Big[\sum_i \cos(\sqrt{4\pi}\phi_i)\Big]^2\right)$$
$$\times \exp\left(i\sqrt{\pi}(\phi_1(R) - \phi_1(0)\right). \tag{103}$$

Consider the square of the sum over cosines. The diagonal terms can be lumped with the free field term using Eq. (83) with two caveats. First, we can ignore the irrelevant term since we will never deal with strong coupling. Second we replace each cosine squared by $-\frac{1}{2\pi}(\nabla\phi)^2$. This is because in the old derivation we brought the two cosines together in the spatial direction, at equal times. Here it is more appropriate to average over the spatial and temporal separations in defining the cosine squared. Imagine now rescaling ϕ so that the kinetic term still has the standard coefficient of $1/2$. In terms of the new field

$$\overline{\phi} = \left(1 + \frac{g^2}{2\pi}\right)\phi \tag{104}$$

(once again called ϕ, in what follows)

$$\overline{G^2(R)} = \left\langle \exp\left(i\sqrt{\frac{\pi}{1+\frac{g^2}{2\pi}}}\phi_1(R)\right)\exp\left(-i\sqrt{\frac{\pi}{1+\frac{g^2}{2\pi}}}\phi_1(0)\right)\right\rangle_g, \tag{105}$$

where the subscript g tells us that the average is taken with respect to the vacuum of an interacting field theory with action

$$S = \int d^2x \sum_1^n -\tfrac{1}{2}(\nabla\phi_i)^2$$
$$+ \frac{g^2}{2\pi^2\alpha^2}\Big[\sum_i \sum_{j\neq i} \cos(\sqrt{4\pi}\phi_i \cos(\sqrt{4\pi}\phi_j)\Big]. \tag{106}$$

Unlike in the pure case, where we had a two point function to evaluate in a free field theory, we have an interacting theory. Since the size of the interaction, g^2, measures the width of the bond distribution, perhaps we can work first with small g in a perturbation expansion? Unfortunately not. If we carry out the series we will find that the real expansion parameter is $g^2 \ln R/a$. Thus no how small g is, we cannot trust anything at large R. There is a well known way out of this problem. Suppose we can trade the theory with cut-off a for a theory with cut-off $2a$. This can be done, say be integrating out every other degree freedom in the functional integral and

using what remains as the Boltzmann weight for the remaining degrees. In the new lattice the spins will be separated by $R/2a$ sites and the log will be a little smaller. But there are two problems. First, the new theory will have many couplings that come from the integration just refered to. Here is where renormalizability comes: if the theory is renormalizable, as ours is, we can lump all the effects of the cut-off change into a suitably chosen g, as long as we ask only questions involving long distances. In addition, the surviving fields in the functional integral will be rescaled by a factor $\mathcal{Z}$ (not to be confused with the partition function) with respect to what we started with. This is because the integration of the unwanted modes will in general rescale the free field part of the action as well. This must be brought to standard form (say $\frac{1}{2}(\nabla\phi)^2$ in the case of a scalar field.) If we do not do this, we cannot call the number in front of the quartic term the coupling constant.

Suppose we do all this. We could still have a problem: the new g is now bigger and offsets that gain we made with the log factor. To see which way things go, one has to compute the β function:

$$\beta(g) = \frac{dg}{d\ln \Lambda}, \tag{107}$$

where Λ is the inverse of the lattice size. This computation involves finding to any given order the contributions the eliminated modes make to the interaction between the surviving modes. To second order in g^2 one draws the three possible one loop graphs and integrated the loop momenta from the old Λ to the new. For the n-component Gross–Neveu model one knows [12] that

$$\beta(g) = (1-n)\frac{g^3}{2\pi} + \text{higher order}. \tag{108}$$

In our case, $n = 0$, the coupling decreases with increasing lattice size. If we increase the lattice size all the way to R, we find, upon integrating the above equation, that

$$g^2(R) = \frac{\pi}{\ln \frac{R}{a}} \tag{109}$$

for very large R/a. Hence, when we do perturbation theory in the new lattice with size R, not only will we be dealing with a small coupling constant $g(R)$, the troublesome logs will be gone.

Consider the two-point function of some operator Ω computed in the original theory with coupling $g(a)$ on a lattice of size a. Call this $G(R, g(a), a)$. We assert that

$$G(R, g(a), a) = \exp\left[2\int_{g}^{g(R)} \frac{\gamma(g)}{\beta(g)} dg\right] G(R, g(R), R). \tag{110}$$

The second factor tells us that as we change the lattice size from a to R, we must change the coupling from $g(a)$ to $g(R)$. The first is just the rescaling referred to earlier. It has this form for the following reason. Imagine doing the lattice change in infinitesimal steps. At each step the coupling that is operative is different and so is the rescaling. Define

$$\gamma(g(\Lambda)) = \frac{dZ}{d\ln\Lambda}. \tag{111}$$

Then the prefactor in Eq. (110) is seen to be just the integral of this equation after two facts are incorporated. First, we need 2 because it is a two-point function, and each of the two operators gets rescaled. Next, the integral over $d\ln\Lambda$ is traded for an integral over g by using the definition of the β function, Eq. (108).

Since G can depend only on the ratio of the distance R to the cut-off, and this ratio equals unity on the in the G in the right hand side of Eq. (110). There is a mild dependence of the G on R via the coupling $g(R)$ contained within. We may evaluate this G at zero coupling, the neglected corrections being down by powers of $g^2 \sim 1/\ln R$. Thus all the leading R dependence is contained in the exponential integral. To do it, we just need the lowest order γ. We obtain this from its definition. First note that if we ignore all interactions (due to the double sum over cosines),

$$\begin{aligned} G(R,g,\Lambda) &= \left\langle \exp\left(i\sqrt{\frac{\pi}{1+\frac{g^2}{2\pi}}}\phi_1(R)\right)\exp\left(-i\sqrt{\frac{\pi}{1+\frac{g^2}{2\pi}}}\phi_1(0)\right)\right\rangle_g \\ &= \left[\frac{1}{R\Lambda}\right]^{\frac{1}{2}\left(1+\frac{g^2}{2\pi}\right)^{-1}} \cdot (1+O(g^4)). \end{aligned} \tag{112}$$

The term in square brackets comes using Eq. (42) valid for the free field theory, its g dependence comes from explicit factors of g in the definition of the operators. The corrections to the interactions begin at order g^4 because the diagonal terms in the double sum have been pulled out and the off-diagonal terms do not contribute to correlation in question due to the constraint that the sum of all the exponents must add up to zero for each boson. Looking at Eq. (112) we can see that if we were to change the cut-off to Λ^* and wanted the same answer the operator to use must get rescaled by a factor

$$Z = \left[\frac{\Lambda}{\Lambda^*}\right]^{\frac{1}{4}\left(1+\frac{g^2}{2\pi}\right)^{-1}}, \tag{113}$$

from which it follows that to order g^2

$$\gamma = \frac{1}{4} - \frac{g^2}{8\pi}. \tag{114}$$

Doing the integral it is easy to obtain

$$\overline{\langle s_R s_0\rangle^2} \sim \frac{(\ln R)^{1/4}}{R^{1/2}}. \tag{115}$$

We now use the fact that the mean of the square is an upper bound on the square of the mean to obtain

$$\overline{\langle s_R s_0\rangle} \leq \frac{(\ln R)^{1/8}}{R^{1/4}}. \tag{116}$$

Thus we find that the DD formula cannot be right since it violates this bound. It is also nice to see the kind of logs you expect in an asymptotically free theory.

Several developments have taken place since this [13] work was done. First I learnt that Shalayev [15] had independently done this, without using bosonization. The English version of the paper was brought to my notice about a year or so later. Next, in my paper I had made to claim that if my arguments were repeated for higher moments one would find that the $2n$-th power of G averaged would be the n-th power of G^2 averaged that I just described. Ludwig pointed out [14] that this was wrong: the error came from using the $\exp(i\sqrt{\pi}\phi)$ in place of $\sin\sqrt{\pi}\phi$. Although the difference due to this is zero so far, it does affect the higher moments. Ludwig in fact carried out the very impressive task of obtaining the full probability distribution for G.

More recently Dotsenko and collaborators [16] did a numerical study to confirm the correctness of my bound and some additional predictions made by Shalayev. For more technical details of my derivation given above see the excellent book by Itzykson and Drouffe [17].

7. Application III: the Hubbard model

We now turn to a problem where the fermion occurs from the beginning instead of after the Jordan–Wigner transformation on spin operators. For a more detailed review if this topic see the article by Emery [18]. Consider nonrelativistic fermions hopping on a lattice in $d = 1$. Let

$$H = -\tfrac{1}{2}\sum_{a,n}\left[\psi_a^\dagger(n)\psi_a(n+1) + \text{h.c.}\right] - \sum_{a,n}\mu\psi_a^\dagger(n)\psi_a(n), \tag{117}$$

where $a = 1, 2$ are two possible spin orientations and μ is the chemical potential. Upon introducing plane wave modes :

$$\psi_a(k) = \sum \psi_a(n)\mathrm{e}^{ikn} \tag{118}$$

we get

$$H = \sum_a \int_{-\pi}^{\pi} (\mu - \cos k)\psi_a^\dagger(k)\psi_a(k)\frac{dk}{2\pi}. \tag{119}$$

In the ground state all fermions with $|k| \leq k_F$ are present, where $\mu = \cos k_F$. The low energy modes of this theory are given by excitations near the Fermi points. As mentioned at the outset, we may linearize the dispersion relations near these two points to obtain a Dirac fermions. Here are some details. First let us break up the k-integral in Eq. (118) to two disjoint parts, centered around $\pm k_F$. Let us keep a band of width $\Lambda \ll k_F$ around each point. The assumption is that for low energy physics these modes will tell the full story and that modes deep in the sea will not be excited. Thus we write

$$\begin{aligned}\psi_a(n) &\simeq e^{-ik_F n}\int_{-k_F-\Lambda}^{-k_F+\Lambda} e^{ikn}\psi_a(k)\frac{dk}{2\pi} + e^{ik_F n}\int_{k_F-\Lambda}^{k_F+\Lambda} e^{ikn}\psi_a(k)\frac{dk}{2\pi} \\ &\equiv e^{-ik_F n}\psi_{a-}(n) + e^{ik_F n}\psi_{a+}(n). \end{aligned} \tag{120}$$

Note that $\psi_{a\pm}(x)$ do not contain high momentum modes. Thus we have written the nonrelativistic field, which does have high momentum modes as explicitly rapidly oscillating factors (the exponential prefactors) times smooth fields. We will assume these fields have slowly varying matrix elements on the scale of the lattice. While this is readily verified in this noninteracting theory (where the correlations fall like inverse distance), it will be assumed to be so even when interactions are turned on.

Feeding Eq. (120) into Eq. (117) we get

$$H_c = \sin k_F \sum_a \int da(\psi_{a-}^\dagger(x)(i\partial_x)\psi_{a-}(x) + (\psi_{a+}^\dagger(x)(-i\partial_x)\psi_{a+}(x)). \tag{121}$$

Those of you wish to derive this must note the following. First, when feed in Eq. (120) there will be cross terms between $\pm$ components. These have rapidly oscillating Fermi factors $e^{\pm ik_F n}$ which will make no contribution. This is because of momentum conservation: the smooth the left and right fields have mean momenta centered around $\pm k_F$ and the two will have no overlap in momentum space since the fluctuations around these means are assumed to be small ($\Lambda \ll k_F$). Next, you must write

$$\psi_{a\pm}(n+1) = \psi_{a\pm}(n) + a\partial_x\psi_{a\pm}n + \cdots.$$

Finally you must remember the relation between μ and k_F, the factor $\sqrt{a}$ which relates the lattice fields to the continuum fields and the factor a which relates H to H_c. (Also feel free to integrate by parts.)

Let us now turn on the Hubbard interaction:

$$H_{\text{int}} = U \sum_n \psi_\uparrow^\dagger(n)\psi_\uparrow(n)\psi_\downarrow^\dagger(n)\psi_\downarrow(n)\,, \tag{122}$$

where ψ stands for the original nonrelativistic fermion. The Hubbard interactions is just the extreme short range version of the screened coulomb potential between fermions. Due to the Pauli principle, only opposite spin electron could occupy the same site. One can extend the model to include nearest neighbor interactions, but we will not do so here.

Let us now express this interaction in terms of the Dirac fields. We get, in obvious notation:

$$\begin{aligned}&\psi_\uparrow^\dagger(n)\psi_\uparrow(n)\psi_\downarrow^\dagger(n)\psi_\downarrow(n) = \\ &(\psi_{\uparrow+}^\dagger(n)\psi_{\uparrow+}(n)+\psi_{\uparrow-}^\dagger(n)\psi_{\uparrow-}(n)+(\psi_{\uparrow+}^\dagger(n)\psi_{\uparrow-}(n)\mathrm{e}^{-2ik_{\mathrm{F}}n}+\text{h.c.}))\bullet(\uparrow\to\downarrow).\end{aligned} \tag{123}$$

If we expand out the products and keep only the parts with no rapidly oscillating factors (momentum conservation) we will, in the generic case, get the following terms:

$$H_{\text{int}} = U(j_{0\uparrow}j_{0\downarrow}) + U(\psi_{\uparrow+}^\dagger(n)\psi_{\uparrow-}(n)\psi_{\downarrow-}^\dagger(n)\psi_{\downarrow+}(n) + \text{h.c.})\,. \tag{124}$$

If we now bosonize these terms as per the dictionary, we get in the continuum (dropping the subscript c for continuum):

$$H = \int dx \tfrac{1}{2}\left[\Pi_\uparrow^2 + (\partial\phi_\uparrow)^2 + \uparrow\to\downarrow\right] + U\left[\frac{\partial\phi_\uparrow\partial\phi_\downarrow}{\pi} + \frac{1}{\pi^2\alpha^2}\cos\sqrt{4\pi}(\phi_\uparrow - \phi_\downarrow)\right]. \tag{125}$$

We can now separate the theory into two parts by introducing charge and spin fields ϕ_c and ϕ_s:

$$\phi_{c/s} = \frac{\phi_\uparrow \pm \phi_\downarrow}{\sqrt{2}}\,. \tag{126}$$

This will give us

$$H = H_c + H_s \tag{127}$$

$$H_c = \int dx \tfrac{1}{2}\left[\Pi_c^2 + (\partial\phi_c)^2 A_c^2\right] \tag{128}$$

$$H_s = \int dx \tfrac{1}{2}\left[\Pi_s^2 + (\partial\phi_s)^2 A_c^2\right] + U\frac{1}{\pi^2\alpha^2}\cos\sqrt{8\pi}\phi_s \tag{129}$$

$$A_{c/s}^2 = 1 \pm \frac{U}{\pi}\,. \tag{130}$$

It is obvious that the charge sector is massless, being a quadratic hamiltonian. This means that there will be no gap to creating charge excitations, the system will be metallic. The fate of the spin sector needs some work. First, we do not like the unconventional free-field term. We cannot simply rescale ϕ (as we did in the functional integral) to bring the coefficient of the gradient term to $1/2$ since it is now an operator, with a canonical commutation relation to satisfy. To preserve the latter, we must scale ϕ and Π oppositely. Let us pull out a factor A_s out of H_s and forget it hereafter since the overall scale of H_s does not concern us. We will now find an inverse A_s in front of Π^2, simply A_s in front of the grad squared and the U in front of the cosine will become $U_s = A_s$. Let us now rescale the momentum and field oppositely by $\sqrt{A_s}$ thereby getting the free-field part to standard form without affect the commutation rules. The argument of the cosine however changes to

$$\text{argument} = \sqrt{\frac{8\pi}{\sqrt{1 - U/\pi}}}\,\phi_s\,. \tag{131}$$

We can now show that for weak positive U, this interaction does not produce any mass, while for weak negative U, it does. Suppose we do perturbative calculation around the free theory with this interaction. At second order we will get two-point correlations of this operator. According to Eq. (42) it will go as

$$G(R) = \frac{1}{\alpha^4}\left[\frac{\alpha^2}{\alpha^2 + R^2}\right]^{\beta^2/4\pi} \tag{132}$$

$$\beta = \sqrt{\frac{8\pi}{\sqrt{1 - U/\pi}}}\,. \tag{133}$$

Thus if $U > 0$, there are more powers of α in the numerator and in the limit of vanishing α, the operator is irrelevent. By the same token for the attractive case, the perturbation process divergent effects, it is relevent. This does not prove that there is a mass, just makes it very likely. Besides the exact solution by Lieb and Wu [19], the following physical argument helps. If there is an attraction between opposite spin electrons, they will tend to form on site, singlet pairs. To make a spin excitation, we must break a pair, and this will cost us, *i.e.*, there will be a gap in the spin sector.

In the special case of half filling, another term comes in. If we look at Eq. (123) we see that in the case of half-filling, since $k_F = \pi/2$, the factors $e^{\pm 4ik_F n}$ are not rapidly oscillating but, simply equal to unity. Thus two previously neglected terms, in which two right movers are destroyed and two left movers are created and vice versa, come to play. (This is an umklapp process, in which lattice momentum is conserved modulo 2π). I

leave it to you to verify that the bosonized form of this interaction, after rescaling of the charge field in the manner described above for the spin field, is another cosine, with

$$\beta = \sqrt{\frac{8\pi}{\sqrt{1+U/\pi}}}. \tag{134}$$

Thus we find the situation is exactly reversed in the charge sector: there is a gap in repulsive case, and no gap in the attractive case. To see what is happening, think of very large positive U. Now there will be one electron per site at half-filling, unable to move without stepping on someone else's toes *i.e.*, there is a charge gap of order U if you try to move the charge. But the spin can do whatever it warns with no cost. If U were very large and negative, there would be tightly bound pairs on half the sites. These doubly charged objects can be moved without cost. There will however be a cost for breaking the spin pair.

With all this formalism in place, we can get another solution for free. Suppose the fermions had no spin. They cannot have any on site interaction since the Pauli principle will not let them occupy the same site. So let us say U is the nearest neighbor coupling. If we proceed as we did above, we will find for generic filling that the theory is massless: you will have to use Eqs. (52) and (54). There will be no cosines in sight to produce a mass, it will all be quadratic terms in grad ϕ. At half-filling, Eq. (54) will admit a term that is irrelevant to begin with. However it is possible that upon sufficient rescaling due to interaction we could end up with $\beta^2 = 8\pi$ though we stared 16π. Indeed, the exact solution by Yang and Yang tells us that this is indeed the case. Please consult references [20] given in the end for more reading on this topic.

8. Conclusions

I hope that with this introduction you can go out and read more things, and more importantly, do more things. I have focused on problems I have worked on since I know these in some detail. You are urged to read some or all of extra references [21] provided. In particular the time is ripe to master what is called nonAbelian bosonization, invented by Witten and very well explained and exploited by Affleck in his review [22] and publications [23].

A more recent development has been the application of bosonization to problems in higher dimensions. The basic idea of Haldane [24] was developed very energetically by the group at Brown University, consisting of Houghton, Kwon and Marston [25].

REFERENCES

[1] D.C. Mattis, E. Lieb, *J. Math. Phys.* **6**, 304 (1965).

[2] A. Luther, I. Peschel, *Phys. Rev.* **B12**, 3908 (1975).

[3] S. Coleman, *Phys. Rev.* **D11**, 2088 (1978).

[4] S. Mandelstam, *Phys. Rev.* **D11**, 3026 (1975).

[5] C. Itzykson, J.B. Zuber, *Phys. Rev.* **D15**, 2875 (1977).

[6] E. Fradkin, L. Susskind, *Phys. Rev.* **D17**, 2637 (1978).

[7] T.D. Schiltz, D. Mattis, E.H. Lieb, *Rev. Mod. Phys.* **36**, 856 (1964).

[8] Vl.S. Dotsenko, V.S. Dotsenko, *Adv. Phys.* **32**, 129 (1983).

[9] See lectures by R. Rajaraman given at the BCSPIN School, Katmandu, May 1991.

[10] See lectures by R. Mani, given at the BCSPIN School, Katmandu, May 1991.

[11] D.J. Gross, A. Nevue, *Phys. Rev.* **D10**, 3235 (1974).

[12] For an answer good to two loops see, W. Wetzel, *Phys. Lett.* **153B**, 297 (1985).

[13] R. Shankar, *Phys. Rev. Lett.* **58**, 2466 (1987).

[14] A.W.W. Ludwig, *Phys. Rev.* **61**, 2388 (1988), see response to this comment R.Shankar *Phys. Rev.* **61**, 2390 (1988).

[15] B.N. Shalayev, *Sov. Phys. (Solid State)* **26**, 1811 (1984).

[16] V.B. Andreichenko, Vl.S. Dotsenko, W. Selke, J.J. Wang, *Nucl. Phys.* **B344**, 531 (1990).

[17] C. Itzykson, J.M. Drouffe, *Statistical Field Theory*, Vol.2, Cambridge University Press, 1990.

[18] V.J. Emery, in *Highly Conducting One-dimensional Solids*, Ed., J.T. Devrese, R.P. Evrard, and V.E. Van Doren, Plenum, 1979.

[19] E.H. Lieb, F.Y. Wu, *Phys. Rev. Lett.* **20**, 1145 (1968).

[20] F.D.M. Haldane *Phys. Rev. Lett.* **45**, 1358 (1968); M.P.M. Den Nijs, *Phys. Rev.* **B23**, 6111 (1981).

[21] For the exact solution to this problem see C.N. Yang, C.P. Yang, *Phys. Rev.* **150**, 321 (1976). Here some more papers that use bosonization. The list is not exhaustive or even representative. However, once you get your hands on these you can follow the leads given therein to find more. T. Giamarchi, H.J. Schulz, *Phys. Rev.***B37**, 325 (1988); S. Sachdev, R. Shankar, *Phys. Rev.* **38**, 826 (1988); R. Shankar, *Phys. Rev. Lett.* **63**, 203 (1989); R. Shankar, *Int. J. Mod. Phys.* **B4**, 2371 (1990). These papers use the same notation and conventions used in these lectures.

[22] I. Affleck in *Fields Strings and Critical Phenomena*, ed. E. Brezin and J. Zinn-Justin, North Holland, 1991.

[23] See for example I. Affleck, A.W.W. Ludwig, *Nucl. Phys.* **B360**, 641 (1991).

[24] *Bosonization in $d = 2$. Luttinger's Theorem and Bosonization on the Fermi Surface, Varenna lectures, 1992,* eds Schriefer and R.Broglia, North Holland 1992.

[25] A. Houghton, H.J. Kwon, J.B. Marston, *Phys. Rev.* **B50**, 1351 (1994).