B.6.1 : SEMI-CLASSICAL METHODS

In the last chapter we already saw how there exist results that are unobtainable in any perturbative expansion about zero coupling. The idea of semiclassical theory is to make a quite different expansion, this time around the classical solution to the theory. Thus what we will end up doing is expanding in the strength of what are usually called the "quantum fluctuations" around the theory.

From a theoretical point of view this makes a lot of sense. In the enormously large "space of solutions" to any QFT, the classical limit is rather far away from the limit of zero coupling. Thus, any non-trivial non-linear solution to the classical field equations is completely inaccessible to a perturbation expansion (except under certain circumstances); and likewise tunneling phenomena (which, as we will see, can be studied using semiclassical theory) are likewise inaccessible.

As we shall see, the expansion around the classical solution is actually an expansion in powers of $\hbar$, and we have actually seen this in ordinary QM, as an expansion in fluctuations around the classical path. It also turns out that in diagrammatic terms, we can imagine that this expansion is an expansion in the number of loops in the relevant graphs - with the classical solution corresponding to a sum of "tree diagrams". However we need to be careful here - it turns out that certain infra-red diagrams containing loops can also contribute to the classical solution.

Of course the classical solutions are not always easy to find, particularly for the more complicated theories in which large varieties of topological and non-topological solitons may exist. For Yang-Mills theories, and for gravity, or for a system like superfluid ^3He, we have only a very small sampling of the huge variety of different possible classical solutions. Moreover for certain systems it is not even clear what is the classical limit of the quantum theory. The simplest example of this is the non-relativistic system of spins in a solid, where the spins are localized on lattice points. Since there is no classical limit for a spin (if we let $\underline{S} = \hbar S \to 0$, we get nothing) this was for a long time a problem, until finally in the 1980's it was found how to write a path integral for spin.

In what follows we will develop the basic theory, and apply it to a few examples. However the emphasis here is on the methods — the number of examples is so large that there is no point in trying to give a general picture of these.

B.6.1(a) THE CLASSICAL LIMIT

Obviously, before we can study fluctuations around the classical limit, we need to understand what this is. Anyone familiar with the enormous complexities of even the simplest classical mechanical system (eg., a set of 3 or more bodies

interacting gravitationally; or the simple vacuum of General Relativity) will appreciate that this may not be a simple matter. To make things more difficult, the limit of a quantum theory as $\hbar \to 0$ is a singular one, so it is by no means clear that the limit is unique, or that by taking it, we actually do end up with the correct classical theory. Here we look at some of these issues.

(i) RELEVANT FEATURES OF CLASSICAL MECHANICS

Let us first recall some results in classical mechanics that inevitably play a role in this discussion. One deals typically with either the Lagrangian L or the Hamiltonian $\mathcal{H}$; let us assume a set of generalized coordinates $\underline{Q} = (q_1 \ldots q_N)$, where N may be infinite, and $\underline{P} = (p_1 \ldots p_N)$; here we are looking at a non-relativistic system to start with. Then we have the action

$$S = \int dt\, L[\underline{Q}, \dot{\underline{Q}}; t) = \int \underline{P} \cdot d\underline{Q} - \mathcal{H}(\underline{P}, \underline{Q})\, dt \tag{1}$$

with, if we minimize the action, the eqtns

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{\underline{Q}}} - \frac{\partial L}{\partial \underline{Q}} = 0 \tag{2}$$

and

$$\left.\begin{aligned} \dot{\underline{Q}} &= \partial \mathcal{H} / \partial \underline{P} \\ \dot{\underline{P}} &= -\partial \mathcal{H} / \partial \underline{Q} \end{aligned}\right\} \tag{3}$$

We also have the differential relations

$$\left.\begin{aligned} \partial S / \partial \underline{Q} &= \underline{P} \\ \frac{\partial S}{\partial t} &= -\mathcal{H}(\underline{P}, \underline{Q}; t) \end{aligned}\right\} \tag{4}$$

and, from (1), that

$$\frac{d}{dt} S = L(\underline{Q}, \dot{\underline{Q}}; t) \tag{5}$$

The 2nd eqtn of (4) is usually written as the Hamilton-Jacobi eqtn., viz.,

$$\frac{\partial S}{\partial t} + \mathcal{H}(\underline{Q}, \partial S / \partial \underline{Q}; t) = 0 \tag{6}$$

and if $\mathcal{H}$ is independent of t, so that $\mathcal{H}_0(\underline{Q}, \partial S / \partial \underline{Q}) = E_0$, we have

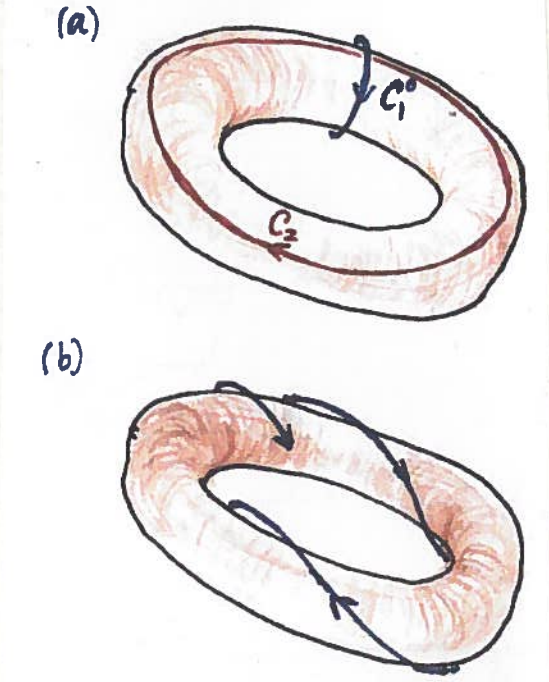
$$S(t) = S_0(\underline{Q}) - E_0 t \tag{7}$$

A key distinction must now be made between integrable and non-integrable systems. Thus we have the following 2 possibilities:

(a) INTEGRABLE SYSTEMS : If a system with $6N$ degrees of freedom (we imagine here a set of N particles moving in 3d space) is completely integrable, it will then possess $3N$ constants of motion. What this means is that in the $6N$-dimensional phase space of the system, it will move on a "$3N$-torus", ie., a $3N$-dimensional hypertorus in the $6N$-dimensional space. There will then be $3N$ natural periods of motion of the system, with periods T_j°, where $j = 1, 2, \ldots 3N$. The motion of the system can then be written in the form

$$\underline{Q}^\circ(t) = \prod_{j=1}^{3N} \sum_{n_j=-\infty}^{\infty} \underline{Q}^\circ_{n_1 n_2 \ldots n_{3N}} \, e^{i \sum_j n_j \omega_j^\circ t} \qquad (8)$$

where $\omega_j^\circ = 2\pi / T_j^\circ$. Each of these $3N$ independent periodic motions corresponds to motion around a closed circuit C_j° around the hypertorus; we show this in the figure for a 2-torus.



(a) TWO INDEPENDENT CLOSED CURVES C_1° AND C_2° ON THE 2-TORUS.

(b) A PATH MADE FROM C_1° AND C_2°, WHICH IS BI-PERIODIC, BUT NOT CLOSED

A simple invariant torus, showing independent orbits, and also a trajectory involving 2 periods that are incommensurate.

It is convenient to rewrite things here in terms of the action-angle variables introduced by Delaunay in 1846. One makes the canonical transformation

$$(\underline{P}, \underline{Q}) \equiv (p_j, q_j) \rightarrow (J_j^\circ, \theta_j^\circ(t)) \qquad (9)$$

where the $\{J_j\}$ are the action accumulated around a circuit C_j°, ie.,

$$J_j^\circ = \frac{1}{2\pi} \oint_{C_j^\circ} \underline{P} \cdot d\underline{Q} \qquad (10)$$

and the θ_j° measure the angle (or "hyperangle") traced out during part of a circuit. Since the canonical relations $p_j = \partial S_0 / \partial q_j$ are invariant under the canonical transformation, we also have $\theta_j^\circ(t) = \partial S_0 / \partial J_j^\circ$, and Hamilton's eqns become

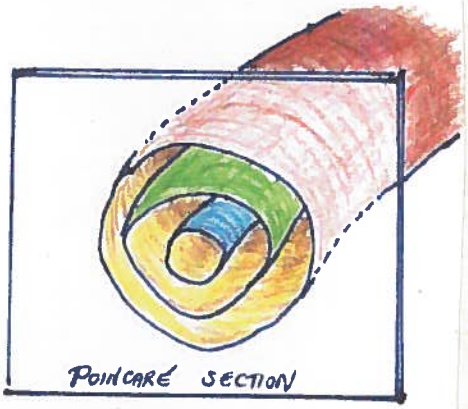
$$\left.\begin{aligned} \dot{J}_j(t) &= -\partial \mathcal{H}_0 / \partial \theta_j^\circ = 0 \\ \dot{\theta}_j^\circ(t) &= \partial \mathcal{H}_0 / \partial J_j^\circ = \omega_j^\circ \end{aligned}\right\} \qquad (11)$$

and so we have

$$\left.\begin{aligned} \omega_j^\circ &= \partial \mathcal{H}_0 / \partial J_j^\circ = \partial E_0 / \partial J_j^\circ \\ \theta_j^\circ(t) &= \omega_j^\circ t + \alpha_j \end{aligned}\right\} \qquad (12)$$

where the $\{\alpha_j\}$ are a set of integration constants. Thus the $\theta_j^\circ(t)$ change linearly with time - the system winds at a constant angular velocity around the circuits C_j° - and ω_j° is a measure of the ratio between the total energy and the action J_j°.

Since J_j° is just a measure of the area of the orbit C_j° in phase space, we see that ω_j° measures how rapidly $E_\circ$ varies as we change this area. Thus we can imagine doing what is shown in the figure, viz., expanding the tori by increasing the energy of the system. As we vary the energy, we get a family of tori, enclosing each other. As Poincaré (the first to seriously investigate all of this) noted, one can analyze the motion over time by making a "Poincaré section" through the hypertorus, and then plotting the points where the system passes through this section on repeated passes (this section is of course a hyperplane for large dimensions). Since the different periods $\{\omega_j^\circ\}$ are not in general commensurable, over time the system trajectory will cover the entire torus (if they are commensurable, it will simply repeat the same trajectory on each pass). Thus on the Poincaré section, for a given fixed energy, we expect to see the trajectory intersections to slowly build a "hypercircle" where the "invariant torus" crosses through the section.



INVARIANT TORI OF DIFFERENT ENERGY, INTERSECTING A POINCARÉ SECTION (FLAT HYPERPLANE).

Now all of this is for an integrable system, with Hamiltonian $\mathcal{H}_\circ$. We can also imagine a system to be partially integrable - this is the more usual case (there will always be some constants of motion, at least for a closed system). But - what happens in this case?

(b) NON-INTEGRABLE SYSTEMS: The famous question first posed over 200 yrs ago by Laplace, partially answered over 100 yrs ago by Poincaré, and given a much more complete answer over 60 yrs ago by "KAM" (Kolmogorov, 1954; Arnold, and Moser, 1962), can be summarized as follows:

Suppose we write the Hamiltonian $\mathcal{H}(\underline{P},\underline{Q}) = \mathcal{H}(\underline{J},\underline{\Theta})$ (where $\underline{\Theta}(t) = \{\Theta_j(t)\}$, and $\underline{J} = \{J_j\}$), in the form

$$\mathcal{H}(\underline{J},\underline{\Theta};t) = \mathcal{H}_\circ(\underline{J}^\circ,\underline{\theta}^\circ) + V(\underline{J},\underline{\Theta};t) \qquad (13)$$

We want to do perturbation theory in $V(\underline{J},\underline{\Theta})$; so let's expand it in terms of the original modes of the system, in the form

$$\left.\begin{aligned} V(\underline{J}_\circ^\circ,\underline{\theta}^\circ;t) &= \prod_{j=1}^{3N} \sum_{n_j} V_{n_1,\ldots n_{3N}}(J_1^\circ, J_2^\circ, \ldots J_{3N}^\circ)\, e^{-i\sum_{n_j} n_j \omega_j^\circ t} \\ &\equiv \sum_{\underline{n}} V_{\underline{n}}(\underline{J}_\circ^\circ)\, e^{-i\underline{n}\cdot\underline{\omega}^\circ t} \end{aligned}\right\} \qquad (14)$$

so that, integrating over time, we have a change in the action of the form

$$\Delta S = i \sum_{\underline{n}}{}' \frac{V_{\underline{n}}(\underline{J}^0)}{\underline{n}\cdot\underline{\omega}^0} e^{-i\underline{n}\cdot\underline{\omega}^0 t} \tag{15}$$

where $\sum'$ indicates we omit the $\underline{n} = 0 \equiv (0,0,\dots)$ term. The change in the solution $\underline{Q}(t)$ is then, in terms of the unperturbed solution $\underline{Q}^0(t)$

$$\Delta \underline{Q}(t) = \int^t dt' \, V(\underline{Q}, t') \, \underline{Q}^0(t') + O(V^2) \tag{16}$$

or, in Fourier components

$$\begin{aligned} \Delta Q(t) &\sim \int^t dt' \sum_{\underline{n}} \sum_{\underline{m}} V_{\underline{n}} Q^0_{\underline{m}} e^{-i(\underline{n}+\underline{m})\cdot\underline{\omega}^0 t'} \\ &\sim i \sum_{\underline{n}}{}' \sum_{\underline{m}}{}' \frac{V_{\underline{n}} Q^0_{\underline{m}}}{(\underline{n}+\underline{m})\cdot\underline{\omega}^0} e^{-i(\underline{n}+\underline{m})\cdot\underline{\omega}^0 t} \\ &\equiv i \prod_{j=1}^{3N} \prod_{k=1}^{3N} \sum_{n_j}{}' \sum_{m_k}{}' \frac{V_{n_1 n_2 \dots n_{3N}} Q^0_{m_1 m_2 \dots m_{3N}}}{n_j \omega_j^0 + m_k \omega_k^0} e^{-i\left(\sum_{n_j} n_j \omega_j^0 + \sum_{m_k} m_k \omega_k^0\right) t} \end{aligned} \tag{17}$$

To see the significance of this result, let us imagine we are dealing with a very simple case in which only 2 frequencies ω_1^0 and ω_2^0 are involved, so that (17) becomes

$$\Delta Q(t) \sim \sum_{n_1 n_2}{}' \sum_{m_1 m_2}{}' \frac{V_{n_1 n_2} Q_{m_1 m_2}}{(n_1+m_1)\omega_1^0 + (n_2+m_2)\omega_2^0} e^{-i[(n_1+m_1)\omega_1^0 + (n_2+m_2)\omega_2^0]t} \tag{18}$$

where now the Σ' excludes vanishing energy denominators – this formula of course breaks down for very long times, when the terms $\sim O(V^2)$ must come in.

Now in general ω_1^0 and ω_2^0 are not commensurate – however we can approximate their ratio ω_1^0/ω_2^0 arbitrarily closely by taking the ratio r_1/r_2 with sufficiently large r_1, r_2, where r_1 and r_2 are integers. Indeed, it is known from number theory that the best approximant $a_{12}(M) = r_1/r_2$, for $\mathrm{Max}(r_1, r_2) = M$, has an error

$$|a_{12}(M)| = \left|\frac{\omega_1^0}{\omega_2^0} - a_{12}(M)\right| \sim O\left(\frac{1}{M^2}\right) \tag{19}$$

so that the error goes down rapidly with increasing M.

This problem of "vanishing energy denominators" was emphasized by Laplace, and Poincaré invented the field of topology in order to try and solve it. We see that as time t increases, the effect of very small energy denominators can become very large – arbitrarily large for high frequency components of the perturbation. Going to higher order in V obviously doesn't

help; we simply introduce chains of energy denominators, and the structure here is one you will recognize from ordinary scattering theory, either in QM (the Lippmann-Schwinger eqtn.) or QFT (the Bethe-Salpeter eqtn.), and can indeed be represented diagrammatically in a very familiar way.

This problem had great interest at that time, in the context of solar system dynamics. Consider, eg., the coupled dynamics of Jupiter and Saturn around the sun. Currently, the periods of these 2 planets are $T_J = 11.8618$ yrs and $T_S = 29.4571$ yrs. The ratio of their orbital frequencies is thus $\omega_S/\omega_J \approx 0.40268$ not so far from 2/5. Their interaction is attractive, and so if they continually meet at the same points in their orbits, there is a strong tendency for this weak attraction to pull them into resonance. In reality ω_S/ω_J fluctuates in time, and Laplace realized that a dynamical system could not only move between different resonant frequency ratios; it could also become unstable. Poincaré uncovered an almost infinitely complex structure in phase space, one that we would now describe as "fractal".

The solution to the problem given by KAM can be stated as

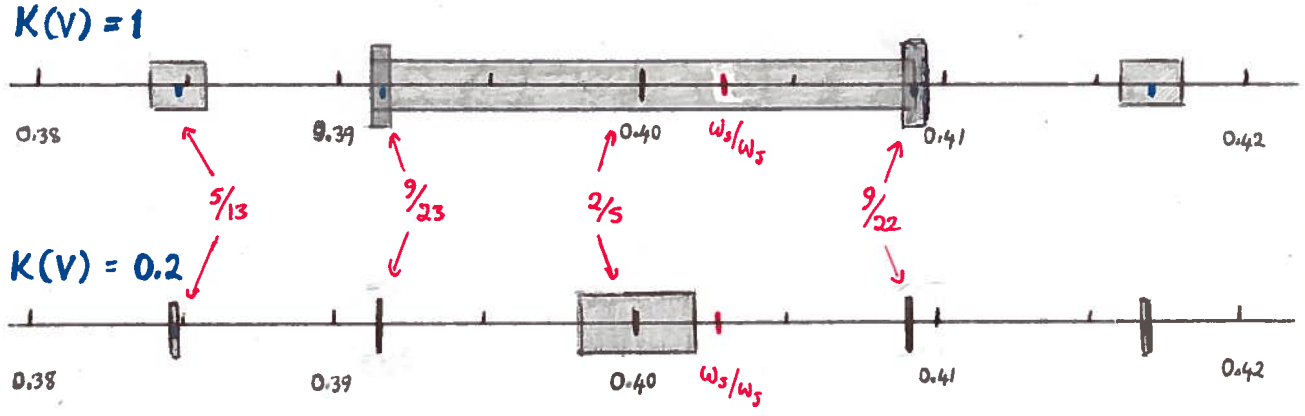
"If $|V|$ is sufficiently small, then for almost all irrational $\underline{\omega}^\circ$ (ie., sets of frequencies $\{\omega_j^\circ\} = (\omega_1^\circ, \omega_2^\circ, \ldots \omega_{3N}^\circ)$ with irrational ratios $\omega_j^\circ/\omega_1^\circ$) there exists an invariant torus $T_V(\underline{\omega}^V)$ of the perturbed system such that $T_V(\underline{\omega}^V)$ is close to the unperturbed torus $T_0(\underline{\omega}^\circ)$."

Moreover, KAM showed that, quite generically, there exists a function of V, usually called $K(V)$, which goes to zero regularly with V, such that the above result is valid provided (here we assume $\omega_2^\circ > \omega_1^\circ$):

$$\left| \frac{\omega_1^\circ}{\omega_2^\circ} - \alpha_{12}(r_2) \right| \equiv \left| \frac{\omega_1^\circ}{\omega_2^\circ} - \frac{r_1}{r_2} \right| > \frac{K(V)}{r_2^{5/2}} \qquad (\forall\ r_2) \qquad (20)$$

in which case we say that the invariant torus is "preserved". It is easy to see that this result leaves most parts of the interval $0 < \alpha_{12} < 1$ preserved, even if V is not that small.

To see how this works out for the Jupiter/Saturn pair, we plot results for (20) for 2 different values of $K(V)$, near the actual ratio ω_S/ω_J.



The region around $\omega_S/\omega_J = 0.40268$ on the line. Regions of size $K(V)/q^{5/2}$ are excluded around any fraction p/q, as shown in shaded rectangles.

We see from this how the KAM theorem works. That only a small fraction of the interval is rendered unstable for small K is seen by making the crude estimate for the measure of the real line that is excised when (20) is not valid; if the measure is $M(K)$, we have

$$M(K) < \sum_{r=1}^{\infty} r \frac{K}{r^{3/2}} = K \sum_{r=1}^{\infty} \frac{1}{r^{3/2}} \sim O(K) \qquad (21)$$

where $M(K)$ is considerably less than the sum on the R.H.S., because this sum double-counts many fractions (thus, eg., $1/2$ is also counted as $2/4, 3/6, 4/8$, etc.).

What happens when the inequality (20) is violated? One possible result is chaotic behaviour, particularly if V is repulsive. Another is the formation of resonantly coupled frequency pairs. We see how complicated is the possible behaviour for just 2 degrees of freedom - for $3N$ coupled degrees of freedom, it is clear that an almost infinitely bewildering variety is possible in the dynamics (including instability).

(c) INTERACTING FIELDS : It is very tempting, although incorrect, to think of the different modes discussed above as oscillators, so that by adding a perturbation we are simply coupling these oscillators. Actually the KAM discussion is much more general - it applies, eg., to orbiting planets, which do not in general behave like oscillators under perturbations.

However a simple interacting bosonic field can be thought of as a set of coupled oscillators. This is clear both for a relativistic scalar field, whose action, when Fourier transformed, reads in the case of ϕ^4 theory as

$$S = -\left\{ \sum_q \tfrac{1}{2} \phi_q (q^2 + m^2) \phi_q + \frac{g}{4!} \sum_{kk'q} \phi_k \phi_{k'} \phi_{k'+q} \phi_{k-q} \right\} \qquad (22)$$

and for a set of interacting phonons (where the coupling is more complicated; see section B.3, eqtn (82)) but of the same general form. Thus one can clearly learn something about coupled fields by studying the simple non-relativistic problem of coupled oscillators, with action

$$S = \int dt \left[\tfrac{1}{2} \sum_q m_q (\dot{x}_q^2 - \omega_q^2 x_q^2) - \sum_{n \geq 3} V_n(q_1 \cdots q_n) x_{q_1} x_{q_2} \cdots x_{q_n} \right] \qquad (23)$$

as an example. The distribution of frequencies $\{\omega_q\}$ is clearly important, but in a typical interacting field, we can assume a quasi-continuous distribution with a density of states $N(\omega)$; in a solid there may exist gaps between different oscillator "bands".

The first serious investigation of a problem like this required a computer (as does much of modern work on non-linear dynamics); it was done by Fermi, Pasta, and Ulam in , in an attempt to understand how a set of oscillators would thermally relax once interactions were introduced. To their great surprise, they found that new modes appeared which did not relax -

we now think of these modes as a kind of "non-topological soliton" which hardly interacts with the other modes.

The situation now is, if anything, more confused. We are far more aware of the enormous variety of possible behaviour in non-linear dynamical systems. Moreover, as we will discuss below, the existence of constraints in the dynamics of a classical system can have unforeseen consequences - this has become clear in the study of solutions to Einstein's eqtns in General Relativity, and is in a sense obvious in subjects like fluid dynamics, and can have important consequences for the quantum theory. However it has to said immediately that there is still very little known about how to give a general theoretical treatment here. As we will see, most work in QFT has focussed on soliton excitations and on their imaginary time counterpart, the "instanton". It is perfectly obvious that this is only scratching the surface of a very large problem.

We will return in detail to all of this later. But it is worthwhile giving a very brief look at one simple example, which is understood in great detail, both in the classical and quantum theories. This is the Sine-Gordon model, which is used both in relativistic and non-relativistic QFT. It takes the following form:

- Relativistic Sine-Gordon Model: This describes a scalar field in $1+1$ dimensions, with action

$$S = \int d^2x \left\{ \tfrac{1}{2} \partial_\mu \phi \partial^\mu \phi + \frac{m^4}{g} \left[\cos \frac{g^{1/2}}{m} \phi - 1 \right] \right\} \tag{24}$$

$$S = \int d^2x \left\{ \tfrac{1}{2} \left(\partial_\mu \phi \partial^\mu \phi - m^2 \phi^2 \right) + \frac{g}{4!} \phi^4 - \frac{1}{6!} \frac{g^2}{m^2} \phi^6 + \cdots \right\} \tag{25}$$

so that for $g/m^2 \ll 1$ this looks like ϕ^4 theory in 2d. The eqtn of motion for $\phi(x)$ is then

$$\partial^2 \phi + \frac{m^3}{g^{1/2}} \sin \frac{g^{1/2}}{m} \phi = 0 \tag{26}$$

which for g/m^2 looks like a free field equation.

- Non-relativistic Sine-Gordon Model: Here we deal with a 1-dimensional system, which contains a periodic potential in the field amplitude, with action:

$$\left.\begin{aligned} S &= \int dt\, dx \left[\tfrac{1}{2} \left(\frac{\partial \Theta(x,t)}{\partial t} \right)^2 - H(x,t) \right] \\ H(x,t) &= \int dt\, dx \left\{ \tfrac{1}{2} c_0^2 \left(\frac{\partial \Theta}{\partial x} \right)^2 + \omega_0^2 \left(1 - \cos \Theta(x) \right) \right\} \end{aligned}\right\} \tag{27}$$

which describes a 1-d field in many different solid-state systems (as well as in polymers like polyacetylene chains). The eqtn. of motion is then

$$(\partial_t^2 - c_0^2 \partial_x^2)\,\theta(x,t) + \omega_0^2 \sin\theta(x,t) = 0 \qquad (28)$$

These 2 systems are identical, and have the same classical physics. It is common to rescale the relativistic form, using

$$\left.\begin{aligned} x &\to \bar{x} = mx \\ t &\to \bar{t} = mt \\ \phi &\to \bar{\phi} = (g^{1/2}/m)\,\phi \end{aligned}\right\} \qquad (29)$$

so that

$$S = \int d^2\bar{x}\ \frac{m^2}{g}\left[\tfrac{1}{2}\bar{\partial}_\mu\bar{\phi}\bar{\partial}^\mu\bar{\phi} + (\cos\bar{\phi} - 1)\right] \qquad (30)$$

with eqtn. of motion:

$$\left(\frac{\partial^2\bar{\phi}}{\partial\bar{t}^2} - \frac{\partial^2\bar{\phi}}{\partial\bar{x}^2}\right) + \sin\bar{\phi}(\bar{x},\bar{t}) = 0 \qquad (31)$$

and conserved energy

$$E = \frac{m^3}{g}\int d\bar{x}\left\{\tfrac{1}{2}\left[\left(\frac{\partial\bar{\phi}}{\partial\bar{t}}\right)^2 + \left(\frac{\partial\bar{\phi}}{\partial\bar{x}}\right)^2\right] + (1-\cos\bar{\phi})\right\} \qquad (32)$$

Thus the rescaled relativistic theory is like the non-relativistic theory, with energy gap $\omega_0 = 1$, and light velocity $c_0 = 1$. They are both Lorentz invariant.

This system has 2 kinds of excitation. There are the small oscillation modes (phonons or photons), with

$$\left.\begin{aligned} \theta_q^{QP}(x,t) &\sim \theta_0 e^{i(qx-\omega_q t)} \\ \omega_q^2 &= \omega_0^2 + c_0^2 q^2 \end{aligned}\right\} \qquad (33)$$

where "QP" signifies "quasiparticle". However there are also "kink" and "anti-kink" solutions, given by

$$\psi_\pm^V(x,t) = 4\tan^{-1}\left[\exp\left(\pm\frac{\omega_0}{c_0}\frac{x-vt}{(1-v^2/c_0^2)^{1/2}}\right)\right] \qquad (34)$$

$$\bar{\Phi}_\pm^V(v;\bar{x},\bar{t}) = \pm 4\tan^{-1}\left[\exp\left(\frac{\bar{x}-v\bar{t}}{(1-v^2)^{1/2}}\right)\right] \qquad (35)$$

in the 2 cases. The existence of these solutions reflects the discrete symmetry present in the system, whereby

$$\left.\begin{aligned} \phi(\bar{x},\bar{t}) &\longleftrightarrow -\phi(\bar{x},\bar{t}) \\ \bar{\phi}(\bar{x},\bar{t}) &\longleftrightarrow \bar{\Phi}(\bar{x},\bar{t}) + 2n\pi \end{aligned}\right\} \qquad (36)$$

with the same for $\theta(x,t)$; we also notice that the Lorentz invariance is explicit in these solutions. A common way to think of these solutions is in terms of the physical model of an infinite chain of pendula hanging in a gravitational field, with nearest neighbours connected by springs. Then the quasiparticles correspond to small oscillations about the energy minimum, and the kinks to configurations in which the line rotates all the way around the upwards vertical, through an angle 2π.

One can also construct exact multi-soliton solutions to this system. However we wish to focus on one simple point, with a view to future developments. Suppose we have a single soliton/kink which is, for simplicity, taken to be stationary ($v = 0$); and we wish to see how its presence affects the quasiparticle modes. This is easy to do; let's assume a solution

$$\theta_{\pm}(x,t) = \psi_{\pm}^{v}(x,t) + \varphi(x,t) \tag{37}$$

where $\varphi(x,t)$ is a very small perturbation around $\psi_{\pm}^{v}(x,t)$. Substituting this into the equation of motion, we get a new equation

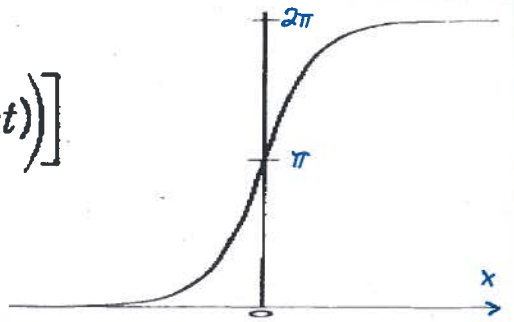
$$\left\{ (\partial_t^2 - c_0^2 \partial_x^2) + \omega_0^2 \left[1 - 2\,\mathrm{sech}^2\left(\frac{\omega_0}{c_0} x\right) \right] \right\} \varphi(x,t) = 0 \tag{38}$$

with the "kink potential" $V(x) = \omega_0^2 [1 - 2\,\mathrm{sech}^2(\omega_0 x / c_0)]$ acting on $\varphi(x,t)$.

$$\psi_{\pm}^{v}(x,t) = 4 \tan^{-1}\left[\exp\left(\pm \frac{\omega_0}{c_0} \gamma (x - vt) \right) \right]$$

$$\gamma \equiv (1 - v^2/c_0^2)^{-1/2}, \quad |v| < c_0$$

The kink soltn for the 1-d Sine-Gordon model.



To solve for $\varphi(x,t)$, we assume the ansatz $\varphi(x,t) = f(x) e^{-i\omega t}$, and substitute in (38). Then we find the eqtn (here $q_0 = \omega_0/c_0$):

$$\left\{ c_0^2 \frac{d^2}{dx^2} + \left[\omega^2 - \omega_0^2 (1 - 2\,\mathrm{sech}^2 q_0 x) \right] \right\} f(x) = 0 \tag{39}$$

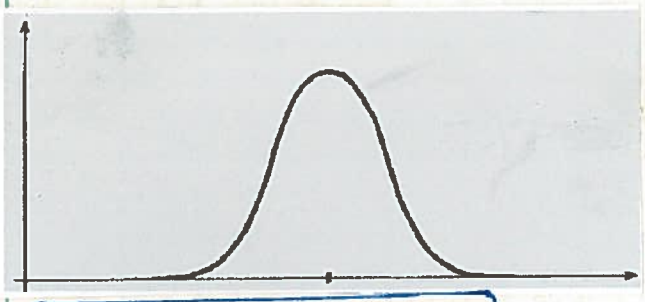
which is a Schrodinger eqtn., with solutions

$$\left.\begin{aligned} f_b(x) &= 2 q_0 \,\mathrm{sech}\, q_0 x \qquad (\text{with } \omega = 0) \\ f_q(x) &= \frac{1}{\sqrt{2\pi}} \left(\frac{c_0}{\omega_q} \right) e^{iqx} \left[q_0 + i q_0 \tanh q_0 x \right] \qquad (\text{with } \omega_q^2 = \omega_0^2 + c_0^2 q^2) \end{aligned}\right\} \tag{40}$$

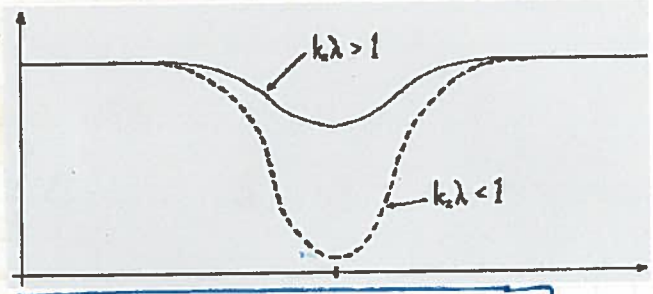
which describe 2 kinds of quasiparticle state. The first, sometimes called a "zero mode", has $\omega = 0$, and is stuck to the soliton as a localized bound state. The second has the form of a running wave, but with reduced amplitude in the region of the kink. This is not surprising - all these modes are orthonormal, so that

$$\left.\begin{aligned} \int dx\, f_q^*(x)\, f_{q'}(x) &= \delta(q-q') \\ \int dx\, f_q(x)\, f_b(x) &= 0 \end{aligned}\right\} \qquad (41)$$

and so we expect the delocalized quasiparticles to have reduced amplitude around the soliton, specially for long wavelengths. All this is illustrated in the figures



$|\psi_b(x)|^2 = f_b^2 = 4q_0^2 \operatorname{sech}^2 q_0 x$



$|\psi_q(x)|^2 = |f_q(x)|^2$ for quasiparticle solns.

One should not imagine that the interaction between the running modes and the kink, or the introduction of a new bound state, means that there is any scattering between the modes. Actually a quasiparticle traveling through the kink suffers a phase shift δ_q given by

$$\delta_q = \pi \operatorname{sign} q - 2\tan^{-1}(q/q_0) \qquad (42)$$

but there is no physical scattering.

(ii) CLASSICAL LIMIT FOR SPIN

: In some cases it is not clear what is the correct classical limit for a quantum system. For any path-integral based method this is a serious problem, because one sums over all paths for a classical action, written in terms of classical variables. This problem emerges most notoriously in the case of spin (a problem which famously defeated Feynman), because in the classical limit $\hbar \to 0$, the spin $\underline{S} = \hbar \underline{\sigma}$, with $|\sigma| = \pm\frac{1}{2}, 1, \ldots$, disappears completely. It has its counterpart in field theories.

In old-style treatments of the problem, one treated spin as though it was an angular momentum. A classical angular momentum $\underline{L}$ has a rotational energy given by

$$T = \tfrac{1}{2}\underline{L}\cdot\underline{\Omega} = \tfrac{1}{2}\Omega_i I_{ij} \Omega_j = \tfrac{1}{2} L_i I_{ij}^{-1} L_j \qquad (43)$$

where $\underline{\Omega} = d\underline{\theta}/dt$ is the angular velocity, and I_{ij} the moment of inertia of the body.

Then the Lagrangian of the system is

$$\mathcal{L} = T - U(\underline{\theta}) \tag{44}$$

where $U(\underline{\theta})$ is a potential energy a function of the orientation angle $\underline{\theta}(t)$ of the object. The Hamiltonian is then $\mathcal{H} = T + U(\underline{\theta})$; and $\underline{L}$ and $\underline{\Omega}$ are related as usual by

$$L_i = I_{ij}\Omega_j \tag{45}$$

Note there is no difficulty in quantizing this and taking the classical limit; if we let $|\underline{L}| = \hbar$, then the classical limit involves letting $\hbar \to 0$ and $\ell \to \infty$ at the same time.

If we try this with spin, we run into a problem. Suppose we try to write the spin as $\underline{S} = \hbar S\,\underline{n}(t)$, where $\underline{n}(t)$ is a unit vector and $S = \frac{1}{2}, 1, \frac{3}{2}, \ldots$ Can we then write a Lagrangian

$$\left.\begin{aligned} L(n,\dot{n}) &= \tfrac{1}{2}\dot{n}_i I_{ij}\dot{n}_j - U(\underline{S}) \\ &\equiv \tfrac{1}{2}\underline{S}\cdot\dot{\underline{n}} - U(\underline{S}) \end{aligned}\right\} \; ? \tag{46}$$

for the spin? If so, we apparently must write $S_i = I_{ij}\dot{n}_j$, but what does this mean? There is no obvious significance to I_{ij}. If we forget about this problem for the moment, one can adopt the following tactic - assume a Hamiltonian $\mathcal{H}(\underline{S})$ for the system, and an equation of motion

$$\dot{\underline{S}}(t) = -\gamma\left(\underline{S}(t) \times \frac{\partial \mathcal{H}(\underline{S})}{\partial \underline{S}}\right) \tag{47}$$

where $\gamma = g\mu_B = \hbar g e/2m$; and then write a Lagrangian

$$L(\underline{S},\dot{\underline{S}}) = P\dot{Q} - H(\underline{S}) \tag{48}$$

with canonical variables

$$\left.\begin{aligned} P(t) &= \hbar S\cos\theta(t) \\ Q(t) &= \varphi(t) \end{aligned}\right\} \tag{49}$$

where the unit vector $\underline{n}(t) = (\theta(t), \varphi(t))$, in polar coordinates. Hamilton's eqtns then give us back (47). While this procedure works, it is also unsatisfactory, because it assumes a special role for the $\hat{z}$-axis, which may not exist in the Hamiltonian.

This problem was solved in a way which is discussed in more detail in part A. One introduces a set of coherent states $|\underline{n}(t)\rangle$, such that

$$\langle \underline{n}|\hat{\underline{S}}|\underline{n}\rangle = \hbar S\,\underline{n} \tag{50}$$

so that, e.g., $\langle \underline{n} | \hat{S}_z | \underline{n} \rangle = \hbar S \cos\theta$, and $\langle \underline{n} | \hat{S}^{\pm} | \underline{n} \rangle = \hbar S \sin\theta\, e^{\pm i\varphi}$. These states thus pick out a specific direction in spin space, or the Bloch sphere, but they are not δ-functions of angle, and two states $|\underline{n}_1\rangle$ and $|\underline{n}_2\rangle$, oriented in different directions, still have a finite overlap. Thus the state $|\underline{n}_z\rangle \equiv |\hat{\underline{z}}\rangle$, oriented along the north pole, we have

$$|\underline{n}_z\rangle \equiv |\hat{\underline{z}}\rangle = |S, m = S\rangle \tag{51}$$

ie., the state with total spin S and projection S along the $\hat{z}$-axis; and for some arbitrary direction $\underline{\theta}$, we have

$$|\underline{n}_{\underline{\theta}}\rangle = |\theta,\varphi\rangle \equiv \hat{R}(\theta,\varphi)\,|\underline{n}_z\rangle \tag{52}$$

where $\hat{R}(\theta,\varphi)$ is just the rotation operator for the vector $\underline{n}$; in terms of the eigenstates $|S,m\rangle$ of $\hat{S}_z$, we have

$$\left.\begin{aligned} |\underline{n}_{\underline{\theta}}\rangle &= \left(\cos\tfrac{\theta}{2}\right)^{2S} \exp\left[\tan\tfrac{\theta}{2}\, e^{i\varphi}\, \hat{S}_-\right] |S,S\rangle \\ &\equiv \sum_{m=-S}^{S} \left(\frac{2S!}{(S-m)!\,(S+m)!}\right)^{1/2} \left(\cos\tfrac{\theta}{2}\right)^{S+m} \left(\sin\tfrac{\theta}{2}\right)^{S-m} e^{im\varphi}\, |S,S\rangle \end{aligned}\right\} \tag{53}$$

and, although these states are complete, ie.,

$$\left.\begin{aligned} (2S+1)\int d\underline{n}\, |\underline{n}\rangle\langle\underline{n}| &\equiv (2S+1)\int \frac{d\theta\, d\varphi}{4\pi} \sin\theta\, |n\rangle\langle n| \\ &= 1 \end{aligned}\right\} \tag{54}$$

we still have the overlap

$$\left.\begin{aligned} \langle \underline{n}_1 | \underline{n}_2 \rangle &= \left[\cos\frac{\theta_1}{2}\cos\frac{\theta_2}{2} + \sin\frac{\theta_1}{2}\sin\frac{\theta_2}{2}\, e^{-i(\varphi_1-\varphi_2)}\right]^{2S} \\ &\equiv \left(\frac{1+\underline{n}_1\cdot\underline{n}_2}{2}\right)^{S} e^{iS\beta_{12}} \end{aligned}\right\} \tag{55}$$

where

$$\tan\frac{\beta_{12}}{2} = \tan\left(\frac{\varphi_1-\varphi_2}{2}\right) \frac{\cos\left(\frac{\theta_1+\theta_2}{2}\right)}{\cos\left(\frac{\theta_1-\theta_2}{2}\right)} \tag{56}$$

Now we can do path integrals - for it turns out - ie., it can be proved that* - the propagator for a single spin between states $|n_1\rangle$ and $|n_2\rangle$ takes the form

$$\langle \underline{n}_2 | \hat{G}(t_2,t_1) | \underline{n}_1 \rangle = \int_{\underline{n}_1(t_1)}^{\underline{n}_2(t_2)} \mathcal{D}n(t)\; e^{\frac{i}{\hbar}\int_{t_1}^{t_2} dt\; L(\underline{n},\dot{\underline{n}};t)} \tag{57}$$

---

* Write $G(2,1) = \lim_{N\to\infty} (2S+1)^N \prod_{j=1}^{N-1} \int d\underline{n}_j \langle \underline{n}_{j+1} | e^{-\frac{i}{\hbar}H(t_j)\,dt} | \underline{n}_j \rangle$, and for infinitesimal dt, rewrite $\langle \underline{n}_{j+1} | e^{-\frac{i}{\hbar}H(t_j)\,dt} | \underline{n}_j \rangle = \langle \underline{n}_{j+1} | \underline{n}_j \rangle \langle \underline{n}_j | e^{-\frac{i}{\hbar}H(t_j)\,dt} | \underline{n}_j \rangle$, and then use (55).

where the Lagrangian is now for a classical vector $\underline{n}(t)$, taking the form

$$L(\underline{n},\dot{\underline{n}};t) = \hbar S\, \underline{A}\cdot\dot{\underline{n}}(t) - \mathcal{H}(S\underline{n};t) \tag{58}$$

with a fake vector potential $\underline{A}$ which is designed to describe the field of a unit monopole situated at the centre of the Bloch sphere, ie., it satisfies

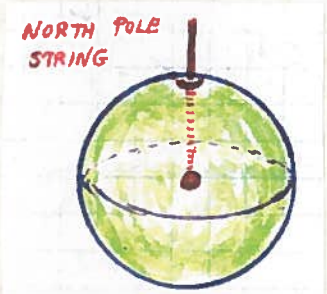
$$\underline{n}\cdot(\underline{\nabla}\times\underline{A}) \equiv n_\alpha \epsilon^{\alpha\beta\gamma}\frac{\partial A^\beta}{\partial n^\gamma} = 1 \tag{59}$$

with the vector operations taken in spin space; in the same way we have

$$\oint_C d\underline{n}\cdot\underline{A} = \omega \tag{60}$$

where the circuit C is taken on the Bloch sphere, and ω is the solid angle enclosed by this circuit.

At this point we have a gauge freedom in choosing $\underline{A}(\theta,\varphi)$, defined as a vector field on the surface of the Bloch sphere. Actually, by the 2-d application of Poincare's conjecture (ie., that every vector field on a sphere has to have a singular point) we know that a "Dirac string" can be used to define this vector field; and in fact the choice of the point where the string emerges is equivalent to a choice of quantization axis for the spin. Let us choose this quantization axis to be the north-directed $\hat{z}$-axis; then one popular choice for $\underline{A}(\theta,\phi)$ is



The "Dirac string" showing the point on the Bloch sphere where $\underline{A}(\underline{n})$ is singular.

$$\underline{A}(\theta,\varphi) = -\hat{\varphi}\cot\theta/2. \tag{61}$$

and another (produced from the first by a non-singular gauge transformation) is

$$\underline{A}(\theta,\varphi) = -\hat{\varphi}\,\frac{1+\cos\theta}{\sin\theta} \tag{62}$$

Thus we have the amusing result that a spin $\underline{S}$, moving in some potential which is a function of the angle $\underline{\theta} = (\theta,\varphi)$, can be described in path integral terms as though it was a particle of charge $\hbar S$, moving on the Bloch sphere surface in the same field, but now with the unit monopole at the centre of the sphere (so that there is a Lorentz force acting on the particle velocity $\dot{\underline{n}}$). This then gives us the correct eqtn of motion (47) for the classical dynamics. However, as will discuss below, it also has profound implications for the quantum dynamics.

This treatment for a single spin is easily adapted to the case of a lattice of spins, to give us a lattice field theory. In all cases the prescription is the same — for a quantum Hamiltonian $\hat{\mathcal{H}}(\hat{\underline{S}})$, written in terms of spin operators, we simply substitute $\hat{\underline{S}} \to \hbar S\underline{n}$, converting the Hamiltonian to a

classical one for the vector $S\underline{n}$. In other words, for a typical exchange + anisotropy Hamiltonian, we have the correspondence

$$\sum_j \hat{\mathcal{H}}_j(\hat{\underline{S}}_j) + \sum_{i\neq j}\sum_j \hat{\mathcal{H}}_{ij}(\hat{\underline{S}}_i, \hat{\underline{S}}_j) \quad \Longleftrightarrow \quad \sum_j H_j(\theta_j,\varphi_j) + \sum_{i\neq j}\sum H_{ij}(\theta_i,\varphi_i;\theta_j,\varphi_j) \tag{63}$$

As an example, consider the Hamiltonian

$$\hat{\mathcal{H}} = \sum_j\left(\frac{1}{S}\left[K_2^{\parallel}\hat{S}_{jz}^2 + K_2^{\perp}S_{jx}^2\right] + \frac{1}{S^3}K_4^{\parallel}S_{jz}^4\right) + \sum_{i\neq j}\sum\frac{1}{S}J_{ij}^{\parallel}\hat{S}_{iz}\hat{S}_{jz} \tag{64}$$

This then has the classical counterpart

$$\left.\begin{aligned} H = {} & \hbar S\sum_j\left(\left[K_2^{\parallel}\cos^2\theta_j + K_2^{\perp}\sin^2\theta_j\cos^2\varphi_j\right] + K_4^{\parallel}\cos^4\theta_j\right) \\ & + \hbar S\sum_{(i\neq j)} J_{ij}\cos\theta_i\cos\theta_j \end{aligned}\right\} \tag{65}$$

with obvious generalizations to other forms of spin Hamiltonian.

If we insist on defining everything in a specific gauge, such as the one employed here in which the Dirac string goes through the north pole, then we can easily recover the results as written in the form (48) and (49). We first need to go to a phase space formulation of the spin path integral, according to which

$$\left.\begin{aligned} G(2,1) = {} & \int_{P_1(t_1)}^{P_2(t_2)}\mathcal{D}P\int_{Q_1(t_1)}^{Q_2(t_2)}\mathcal{D}Q\, e^{\frac{i}{\hbar}S[P,Q]} \\ S[P,q] = {} & \int_{t_1}^{t_2}dt\,\left(P\dot{Q} - H[P,Q]\right) \end{aligned}\right\} \tag{66}$$

However, we do not pursue this here.

B.6.1 (6) FLUCTUATIONS & THE LOOP EXPANSION

Now we return to the quantum theory, and look at expansions around the classical result in the functional integral formulation. This kind of expansion has its roots in simple WKB theory, which we begin by briefly recalling here. Suppose we want to solve a differential eqtn containing a very small parameter, such as the Schrodinger eqtn, written as

$$[\hbar^2 - f(x)]\psi(x) = 0 \tag{67}$$

where, for a 1-particle system, $f(x) = 2m(V(x) - E)$. The basic idea of WKB theory is to write the solution in the form

$$\psi(x) = A(x)\,e^{\Phi(x)/\hbar} \equiv e^{\frac{1}{\hbar}\sum_{n=0}^{\infty}\hbar^n\varphi_n(x)} \tag{68}$$

where $A(x)$ is assumed to very slowly, compared to $\Phi(x)$, in the limit where

$\hbar \ll 1$. Note that this limit $\hbar \to 0$ is singular, as we can see by looking at any simple QM problem; for example, take the "particle in a box", for a specific value of $2m(V-E) = -2mE$ (ie., for specific finite energy). As $\hbar \to 0$, we go to ever higher energy levels for the same energy, and the solution becomes more and more rapidly oscillatory - the limit $\hbar = 0$ is discontinuous, because for $\hbar = 0$, the solution to (67) is $\psi(x) = 0$. Thus (68) is obviously an asymptotic series. Substituting (68) into (67) gives the hierarchy

$$\left.\begin{aligned} (\varphi_0'(x))^2 &= f(x) \\ 2\varphi_0'(x)\varphi_1'(x) + \varphi_0''(x) &= 0 \\ 2\varphi_0'(x)\varphi_n'(x) + \varphi_{n-1}''(x) + \sum_{r=1}^{n-1} \varphi_r' \varphi_{n-r}'(x) &= 0 \qquad n \geq 2 \end{aligned}\right\} \quad (69)$$

a result which we will use several times in this chapter. The solutions for the $\varphi_n(x)$ then have the following results

- The lowest term $\varphi_0(x)$ is just the classical action, ie., we have the general result

$$\varphi_0(x) \equiv S_{cl}(x) = \pm \int^x dx' \, f^{1/2}(x) \quad (70)$$

where in the case $f(x) < 0$, this becomes imaginary.

- The next term is the leading contribution to the prefactor $A(x)$ in the usual QM formula, ie., we have

$$\varphi_1(x) \equiv A_0(x) = -1/4 \ln f(x) \quad (71)$$

corresponding in path integral theory to the 1st fluctuation determinant (see below).

- The next set of terms has the following form, and corresponds, in the path integral theory of QM, to the higher fluctuations about the classical path:

$$\varphi_2(x) = \pm \int dx' \left[\frac{1}{8} \frac{f''(x')}{f^{3/2}(x')} - \frac{5}{32} \frac{(f'(x'))^2}{f^{5/2}(x')} \right] \quad (72)$$

$$\varphi_3(x) = \left[\frac{5}{64} \frac{(f'(x))^2}{f^3(x)} - \frac{1}{16} \frac{f''(x)}{f^2(x)} \right] \quad (73)$$

and so on; for a more detailed discussion one should go to part A.

These results are reflected in the path integral formulation, although to get them requires the calculation of higher fluctuation determinants - all of this is

part of our discussion of path integrals in part A.

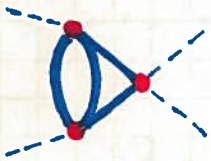
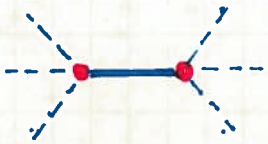
One other thing should be noted here, viz., the connection between this expansion in powers of $\hbar$ and a "gradient expansion", ie., an expansion in higher derivatives of both the function $f(x)$ and the solutions $\varphi_n(x)$ which feed into $\psi(x)$. We see that higher powers in $\hbar$ are associated with higher derivatives (or powers of derivatives) of $f(x)$.

(i) LOOP EXPANSION for SCALAR FIELDS :

We now look at the connection between the asymptotic expansion in powers of $\hbar$ and the "loop expansion", and the usual scalar field theory provides a nice example of this. The basic point here is quite simple, and we shall look at it in 2 different ways: first, by looking at things purely diagrammatically, and then by looking at things as a fluctuation expansion about a classical configuration.

DIAGRAMMATIC ARGUMENT :

Let's consider how the $\hbar$ factors enter the diagrams in a theory like ϕ^4 theory. Let's begin with a table showing how all this works for the lowest-order graphs for the irreducible vertex $\Gamma_n(x_1 \cdots x_n)$ (ie., having no external legs). We work in 4-momentum space, and find:

GRAPH	EXPRESSION	$O(\hbar)$	L
	$\left(\frac{-ig}{\hbar}\right)^3 \sum_{kQ} \sum (i\hbar)^4 \Delta_F(Q)\Delta_F(Q+k-p) \times \Delta_F(k)\Delta_F(p+p'-k)$	$\hbar$	2
	$\left(\frac{-ig}{\hbar}\right)^2 i\hbar\,\Delta_F(k)$	$\hbar^{-1}$	0
	$\frac{-ig}{\hbar} \sum_k i\hbar\,\Delta_F(k)$	$\hbar^0$	1
	$\left(\frac{-ig}{\hbar}\right)^2 \sum_{kq} \sum (i\hbar)^3 \Delta_F(k+q)\Delta_F(k)\Delta_F(p-q)$	$\hbar$	2
	$\left(\frac{-ig}{\hbar}\right)^3 \sum_{kk'q} \sum (i\hbar)^5 \mathcal{F}(k,k',p,q)$	$\hbar^2$	3

with $\mathcal{F}(k,k'pq) = \Delta_F(p-q)\Delta_F(k+q)\Delta_F(k) \times \Delta_F(k'+q)\Delta_F(k')$

where as usual we have the rules given in section B.3 (eqtns (40) and (41) in the Figures) for the vertex and the propagator.

The result is clear already from looking at these graphs; we have, in any graph, a number of loops given by eqtn. (43) of section B.3, ie., that $L = I + 1 - V$, where there are I internal lines and V vertices; and the graph will obey

$$\Gamma_n(k_1..k_n) = \hbar^{L-1} f_n(k_1..k_n) \qquad (74)$$

where the $\{k_n\}$ are the external leg arguments, and $f_n(k_1..k_n)$ is independent of $\hbar$. To prove this result is simple; each internal line gives a factor $\hbar$, while each vertex gives a factor $\hbar^{-1}$; thus we get a contribution $\hbar^{I-V}$ from any graph, ie., $\hbar^{L-1}$.

From this result we see that by summing all "tree graphs" (containing no loops at all) for the vertex generating functional, we will get the complete classical result for the classical generating functional $\Gamma_{cl}[\bar\Phi]$ in terms of the classical solution $\bar\Phi(x)$ of the field theory. We have already seen this result in section B3; cf eqtn (37). There we showed it by expanding $\Gamma_{cl}[\bar\Phi]$ in a graphical expansion, and noting that we got tree graphs; here we have proved it by a general analysis of graphs for the theory.

Actually this proof is not watertight - it was discovered in 2004 that in the case where there are massless fields in the theory (eg., in QED) it is possible for the classical and quantum contributions to mix together at the same order in a loop expansion.* This arises when there are non-analytic terms in the internal lines which yield square-root infrared singularities. - which happens when there are two or more massless propagators in the diagram. This topic is outside the scope of these lectures.

FUNCTIONAL INTEGRAL ARGUMENT: In the path integral formulation of QM, we have seen in part A how one can extract the leading correction to the classical contribution to the 1-particle propagator, in the form of a prefactor or "fluctuation determinant. What we wish to do here is see how this works for a theory like ϕ^4 theory; we also want to see the structure of the whole asymptotic expansion, not just the first correction.

As a warm-up, let's do the problem for ordinary QM, but now going beyond the 1st correction. Thus, we consider the amplitude

$$\left.\begin{aligned} G_0(2,1) &= \int_1^2 \mathcal{D}q(t)\, e^{\frac{i}{\hbar}S[q]} \\ &= \int_1^2 \mathcal{D}q(t)\, e^{\frac{i}{\hbar}\int_1^2 dt\,[\frac{1}{2}m(\dot q^2 - \omega^2 q^2) - V(q)]} \end{aligned}\right\} \qquad (75)$$

where $V(q)$ includes the non-quadratic part of the potential, and the coordinate q is 1-d.

---

* See, eg., B.R. Holstein, J.F. Donoghue, Phys. Rev. Lett. 93, 201602 (2004); and S.J. Brodsky, P. Hoyer, Phys. Rev. D83, 045026 (2011).

Then, in a simple generalization of the usual procedure for extracting the fluctuation determinant, we write

$$q(t) = Q_0(t) + x(t) \tag{76}$$

where $Q_0(t)$ is the classical trajectory, and $x(t)$ the deviation from this trajectory; We then have

$$G(2,1) = A(2,1)\, e^{\frac{i}{\hbar} S_{cl}[Q_0]} \tag{77}$$

where $S_{cl}[Q_0]$ is the classical action, and the prefactor $A(2,1)$ is

$$A(2,1) = \int_{\bar{x}(t_1)=0}^{\bar{x}(t_2)=0} \mathcal{D}\bar{x}(t)\, \exp\left\{-i \int_{t_1}^{t_2} dt \left[\frac{m}{2}\bar{x}(t)\left(\frac{d^2}{dt^2} + \omega_0^2\right)\bar{x}(t) + \sum_{n=3}^{\infty} \frac{\alpha_n}{n!}\, \bar{x}^n(t)\right]\right\} \tag{78}$$

where the coefficient α_n is

$$\alpha_n = \hbar^{\frac{n}{2}-1} \left.\frac{d^n V(\bar{x})}{d\bar{x}^n}\right|_{\bar{x}=0} \tag{79}$$

so that we have connected the higher terms with higher powers of $\hbar$; since only terms of even order contribute, this becomes an expansion in integer powers of $\hbar$ (NB: rescaled $x(t)$ to $\bar{x}(t) = x(t)/\hbar^{1/2}$).

If we ignore the higher terms in $\hbar$, including only the leading correction to the classical result, then we find

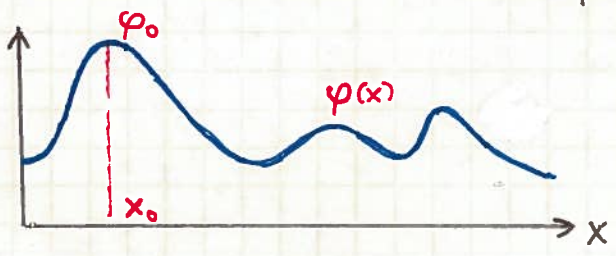
$$\left.\begin{aligned} A(2,1) &= e^{-\frac{1}{2} \ln \det \left| m\left(\frac{d^2}{dt^2} + \omega_0^2\right)\right|} \\ &= \left(\det \left| m\left(\frac{d^2}{dt^2} + \omega_0^2\right)\right|\right)^{-1/2}. \end{aligned}\right\} \tag{80}$$

where as usual the determinant refers to the product of eigenvalues of the differential operator (later on we will rewrite this determinant in terms of a generalized zeta function).

In what follows I will not calculate in full the form of the higher corrections, but simply indicate their form. This can be understood by considering the asymptotic expansion of the integral

$$I(\hbar) = \int_{-\infty}^{\infty} dx\, e^{-\varphi(x)/\hbar} \tag{81}$$

where we assume that the function $\varphi(x)$ has an absolute maximum at a value $x = x_0$, given by $\varphi(x_0) = \varphi_0$. Writing $z = x - x_0$, we then make the expansion



$$\varphi(z) = \varphi_0 + \sum_{n=2}^{\infty} \varphi_n z^n \tag{82}$$

using Laplace's method (which is just a special case of the method of steepest descents); this would be produced from (78) by rotating to Euclidean space, and specializing to a 1-d integral). Then standard methods give us the result

$$\left.\begin{aligned} I(\hbar) &= e^{-\varphi_0/\hbar} \int_{-\infty}^{\infty} dz\, e^{-\frac{\varphi_2}{2\hbar} z^2}\, e^{-\frac{1}{\hbar}\sum_{n=3}^{\infty} \frac{\varphi_n}{n!} z^n} \\ &= \left(\frac{2\pi\hbar}{\varphi_2}\right)^{1/2} e^{-\varphi_0/\hbar} \int_{-\infty}^{\infty} \frac{d\bar{z}}{\sqrt{2\pi}}\, e^{-\bar{z}^2/2}\, e^{-\sum_{n=3}^{\infty} \frac{1}{n!} \frac{\varphi_n}{(\varphi_2)^{n/2}} \hbar^{\frac{n}{2}-1} \bar{z}^n} \end{aligned}\right\} \quad (83)$$

which when expanded out gives

$$\ln I(\hbar) = -\left\{ \varphi_0/\hbar + \tfrac{1}{2} \ln\left(\frac{\varphi_2}{2\pi\hbar}\right) + \frac{\hbar}{\varphi_2}\left[\frac{1}{8}\frac{\varphi_4}{\varphi_2} - \frac{5}{24}\left(\frac{\varphi_3}{\varphi_2}\right)^2\right] + O(\hbar^2) \right\} \quad (84)$$

which for a path integral is generalized to infinite dimensions.

Let us now return to ϕ^4 theory. The classical solution was already discussed in section B.3 ; again making a Euclidean rotation, we have the classical solution $\varphi_0(x)$ satisfying the eqtn. of motion

$$(\partial^2 - m^2)\,\Phi_0(x) - \frac{g}{3!}\Phi_0^3(x) + J(x) = 0 \quad (85)$$

(cf. section B.3, eqtn (29) ; we write $\Phi_0(x)$ here in place of the $\bar{\phi}(x)$ used there).

Now making the expansion in the deviation $\psi(x) = \phi(x) - \Phi_0(x)$, we have

$$\left.\begin{aligned} S[\phi, J] = S_{cl}[\Phi_0, J] + \int d^4x \Big\{ &\tfrac{1}{2}\left[(\partial\psi)^2 + \left(m^2 + \frac{g\Phi_0}{2}\right)\psi^2\right] \\ &+ \frac{g}{3!}\Phi_0\,\psi^3 + \frac{g}{4!}\psi^4 \Big\} \end{aligned}\right\} \quad (86)$$

where the terms linear in $\psi(x)$ vanish because of the eqtn of motion (85). We now expand the action in powers of $\hbar$, in exactly the same way as we have done above. The result can be written as

$$e^{-S[\phi,J]/\hbar} = N^{-1} e^{-S_{cl}[\Phi_0,J]} \int \mathcal{D}\psi\, e^{-\Delta S[\psi]} \quad (87)$$

where the classical action is

$$S_{cl}[\Phi_0, J] = \int d^4x \left[J(x)\Phi_0(x) - \tfrac{1}{2}\left((\partial\Phi_0)^2 + m^2\Phi_0^2\right) - \frac{g}{4!}\Phi_0^4 \right] \quad (88)$$

in terms of the solution $\Phi_0(x)$ to the classical eqtn of motion (85), with the non-zero source field $J(x)$ included; where the normalization factor is

$$N = \int \mathcal{D}\psi \; e^{-\frac{1}{\hbar}\int d^4x \left[\frac{1}{2}\left((\partial\psi)^2 + m^2\psi^2\right) + \frac{g}{4!}\psi^4 \right]} \tag{89}$$

and the shift in the action is, from (86)

$$\Delta S[\psi] = \frac{1}{\hbar}\int d^4x \left\{ \frac{1}{2}\left[(\partial\psi)^2 + \left(m^2 + g\frac{\Phi_0}{2}\right)\psi^2\right] + \frac{g}{3!}\Phi_0\psi^3 + \frac{g}{4!}\psi^4 \right\} \tag{90}$$

Making the same rescaling, viz., $\bar{\psi} = \hbar^{1/2}\psi$, the expansion in powers of $\hbar$ becomes explicit; we get

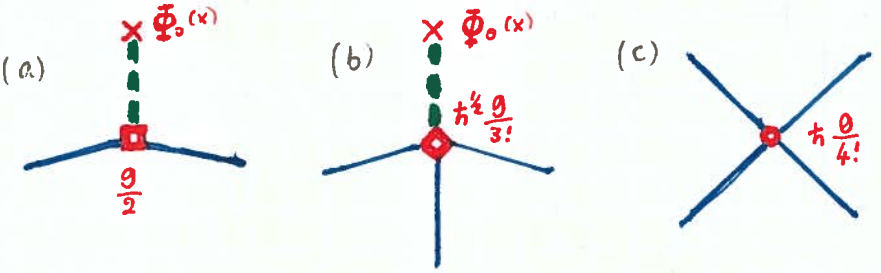
$$\left.\begin{aligned} \Delta S[\psi] &= \int d^4x \left\{ \frac{1}{2}\left[(\partial\psi)^2 + \left(m^2 + g\frac{\Phi_0}{2}\right)\psi^2\right] + \hbar^{1/2}\frac{g}{3!}\Phi_0\psi^3 + \hbar\frac{g}{4!}\psi^4 \right\} \\ N &= \int \mathcal{D}\psi \; \exp\left\{ -\int d^4x \left[\frac{1}{2}\left((\partial\psi)^2 + m^2\psi^2\right) + \hbar\frac{g}{4!}\psi^4 \right] \right\} \end{aligned}\right\} \tag{91}$$

The 1st correction to the classical action is obtained by what we now familiar methods, and we get the result we already know (eqtns (107) and (108) of section B.3), by doing the Gaussian integral. Thus we get

$$S[\phi, J] = S_{cl}[\Phi_0, J] + \frac{\hbar}{2}\ln\left[\frac{\det\left(\partial^2 + m^2 + \frac{1}{2}g\Phi_0^2\right)}{\det\left(\partial^2 + m^2\right)}\right] + O(\hbar^2) \tag{92}$$

so that the role of the "frozen field" $V_0(x)$ in section B.3 is now played by the classical field $\Phi_0(x)$; otherwise everything we did from eqtns (109) - (114) in section B.3 can be applied directly here. Note in particular the x-dependent effective mass, so that $m^2 \to m^2 + \frac{1}{2}g\Phi_0(x)$.

There are however some key differences, notably in the interaction terms in $\Delta S[\psi]$; we have a new coupling $\propto \psi^3$, and in fact the theory boasts 3 different interaction vertices, as shown in the figure.



(93)

The rules for the number of loops and vertices are different here; if V_3 is the number of vertices involving ψ^3, and V_4 those involving ψ^4, we find that the number of loops is

$$L = 1 + \frac{1}{2}V_3 + V_4 \tag{94}$$

Now all of this is only as good as our classical solution $\Phi_0(x)$, since this enters as the "background field". If we know $\Phi_0(x)$, then the

problem becomes one of a self-interacting "quantum fluctuation field" $\psi(x)$, coupled also to a dynamical perturbation $\phi_0(x)$, which affects its mass and adds a new vertex $\propto \psi^3$.

Unfortunately it is hard to find $\Phi_0(x)$; its eqtn of motion (85) is non-linear, and we have seen how difficult such problems are, even for a few interacting particles. Suppose, however, we could find it - what then? Then the fluctuations $\psi(x)$ can, at first approximation, by dropping the interaction terms, since they are higher-order in $\hbar$. Then, with only the vertex (3) in (93) left, we can try solving for the problem - the results being encapsulated in (92).

More generally, we can ask - how good is this semiclassical starting point? This is a hard question to answer in relativistic QFT, since we can't easily modify the fields or the couplings. However one can study toy models, and/or use numerical approaches.

In condensed matter systems it is a quite different story. One can modify the couplings (using external fields, pressure, etc.). And we can even modify the fields: for example, in the FQHE (Fractional Quantum Hall Effect) one modifies the underlying quantum fields by changing the external magnetic field, and in spin systems, one can just change the spin. The key message that emerges is that in many condensed matter systems, the real behaviour can be very far from classical. In other words, the quantum field is very different from the classical field.

B.6.1 (c) QUANTUM SPIN FLUCTUATIONS

Spin systems teach us many interesting lessons — we shall see some here, and more later on. One can write down a non-relativistic field theory for various spin systems — some of these turn out to be closely related to important relativistic QFT's. Remarkably, given that spin has no classical limit, some of these theories are very well described by the classical limit (eg., ferromagnetic systems); others not at all (eg., 1-dimensional antiferromagnetic systems).

Here we will look at both of these cases, and also at how the theory relates to both the real experimental world and to relativistic QFT. The functional formulation turns out to be very powerful, but other methods are also useful. This is a huge subject, and we will only scratch the surface here.

(i) GENERAL REMARKS : Before beginning, it is useful to say a little more about the functional formulation of spin dynamics, and how one can represent fluctuations about the classical state.

SINGLE SPIN DYNAMICS : We start with a single spin, where as we saw in eqtn. (57), the spin propagator can be written as a path integral; between 2 arbitrary states we have

$$\langle \psi_2 | \hat{G}(t_2,t_1) | \psi_1 \rangle = \int d\underline{n}_1 \int d\underline{n}_2 \langle \psi_2 | \underline{n}_2 \rangle \langle \underline{n}_2 | \hat{G}(t_2,t_1) | \underline{n}_1 \rangle \langle \underline{n}_1 | \psi_1 \rangle \qquad (95)$$

with $\langle \underline{n}_2 | \hat{G}(t_2,t_1) | \underline{n}_1 \rangle$ given in (57). So then - what is the classical limit here?

Let's begin by noting that the first "kinetic" term in the Lagrangian (58) is proportional to $\hbar$. Then it follows that

$$\left.\begin{aligned} G(\underline{n}_2,\underline{n}_1;t_2t_1) &= \int_1^2 \mathcal{D}n(t)\, e^{iS\int_1^2 d\underline{n}\cdot\underline{A}}\, e^{-\frac{1}{2}\int_1^2 dt\, \mathcal{H}(S\underline{n})} \\ &\equiv \int_1^2 \mathcal{D}n(t)\, e^{iS\omega_{21}[\underline{n}]}\, e^{-\frac{1}{2}\int_1^2 dt\, \mathcal{H}(S\underline{n})} \end{aligned}\right\} \quad (96)$$

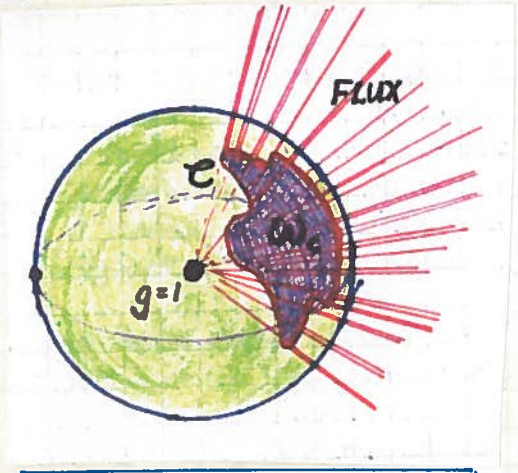
where the number $\omega_{21}[\underline{n}]$ is defined as

$$\omega_{21}[\underline{n}] = \int_{\underline{n}_1}^{\underline{n}_2} d\underline{n}(t)\cdot\underline{A} \quad (97)$$

and is just the contribution from the field of the monopole in the Bloch sphere - the phase is thus completely analogous to the Aharonov-Bohm phase we saw in Chapter 4. To make the analogy complete, let's imagine that the vector $\underline{n}(t)$ accomplishes a circuit, ie., that the spin rotates along some path and then comes back to where it started. One simple way to do this is by making the spin precess, ie., to apply a constant field $\underline{B}_0$ to the spin, so that the Hamiltonian is

$$\left.\begin{aligned} \mathcal{H} &= -\gamma\, \underline{B}_0\cdot\underline{S} \\ &\equiv -\gamma S\, \underline{B}_0\cdot\underline{n} \end{aligned}\right\} \quad (98)$$

which is a simple problem to analyze; but there are other more interesting things one can do, as we will see in the next sub-section.



The Bloch sphere, with monopole at its centre; the enclosed flux by a curve $\mathcal{C}$ is the solid angle Ω_C

If we allow the spin to accomplish a complete circuit $\mathcal{C}$, then the gauge-dependent and path-dependent number $\omega_{21}[\underline{n}]$ now becomes a topological phase (using a suggestive notation for an integration over circuits):

$$\left.\begin{aligned} G(\underline{n},\underline{n};t) &= \oint_{\underline{n}(0)}^{\underline{n}(t)} \mathcal{D}n'(t)\, e^{iS\omega_{\mathcal{C}}[\underline{n}']}\, e^{-\frac{1}{2}\int_0^t dt\, \mathcal{H}(S\underline{n}')} \\ &\equiv \oint \mathcal{D}\mathcal{C}(\tau)\, e^{iS\omega_C}\, e^{-\frac{1}{2}\int d\tau\, \mathcal{H}(S\underline{n}(\mathcal{C}(\tau)))} \end{aligned}\right\} \quad (99)$$

where the phase is just a Berry phase Φ_B, ie.

$$\Phi_B = S\omega_C \quad (100)$$

and we can think of Ω_C, which is just the solid angle enclosed by the circuit C, as a measure of the flux through the circuit (compare eqtn. (60)). Then we have the precise analogy with the Aharonov-Bohm effect (compare eqtns (2) and (36) in chapter B.4).

Now to get a classical limit for this, we need to take a large spin, ie., we need to minimize the action

$$S[\underline{n}] = \int_{t_1}^{t_2} dt \left[\hbar S \underline{A} \cdot \dot{\underline{n}}(t) - \mathcal{H}(S\underline{n}) \right] \tag{101}$$

in the limit $S \gg 1$. Thus we are looking for a path $\underline{n}_0(t)$ on the Bloch sphere which is the solution to

$$\delta S[\underline{n}, \dot{\underline{n}}]\Big|_{\underline{n} \to \underline{n}_0} = 0 \tag{102}$$

Expanding out (100), we find:

$$\delta S = \hbar S \int_{t_1}^{t_2} d\tau \left\{ \left[\left(\frac{\partial A^\alpha}{\partial n^\beta} \delta n^\beta \right) \dot{n}^\alpha + A^\alpha \frac{d}{dt} \delta n^\alpha \right] \Big|_{\underline{n} = \underline{n}_0} - \frac{\partial H}{\partial n^\alpha} \delta n^\alpha \Big|_{\underline{n} = \underline{n}_0} \right\} \tag{103}$$

and the first topological term in this can be rewritten, adding and subtracting a term $(\partial A^\alpha / \partial n^\beta) \dot{n}_\beta \delta n^\alpha$, to give

$$\begin{aligned} \delta S_{Top} &= \delta \int_{t_1}^{t_2} d\tau \, \hbar S \underline{A} \cdot \dot{\underline{n}}(\tau) \Big|_{\underline{n} = \underline{n}_0} \\ &= \hbar S \int_{t_1}^{t_2} d\tau \left\{ \frac{\partial A^\alpha}{\partial n^\beta} \left[\dot{n}^\alpha \delta n^\beta - \dot{n}_\beta \delta n^\alpha \right] \Big|_{\underline{n} = \underline{n}_0} + \left[A^\alpha \frac{d}{dt} \delta n^\alpha + \frac{\partial A^\alpha}{\partial n^\beta} \dot{n}_\beta \delta n_\alpha \right] \right\} \\ &= \hbar S \int_{t_2}^{t_1} d\tau \left[\epsilon_{\alpha\beta\gamma} \frac{\partial A^\alpha}{\partial n_\beta} (\dot{\underline{n}} \times \delta \underline{n})^\gamma + \frac{d}{dt} (A^\alpha \delta n_\alpha) \right] \Big|_{\underline{n} = \underline{n}_0} \end{aligned} \tag{104}$$

Now the 2nd term in this last result is zero, since $\delta \underline{n} = 0$ at the end-points; and the 1st term is rewritten using (59) to give

$$\delta S = \int_{t_1}^{t_2} d\tau \left[\hbar S \epsilon_{\alpha\beta\gamma} n^\alpha \dot{n}^\beta - \frac{\partial \mathcal{H}}{\partial n^\gamma} \right] \Big|_{\underline{n} = \underline{n}_0} \delta n^\gamma \tag{105}$$

so that we get the equation of motion for $\underline{S}_{cl} = \hbar S \underline{n}_0$ from (102) as

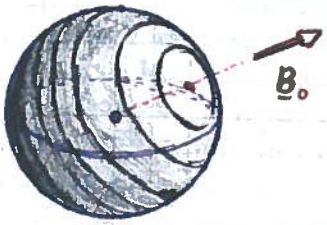
$$\dot{\underline{S}}_{cl} \equiv \underline{n}_0 \times \frac{\partial \mathcal{H}}{\partial \underline{S}_{cl}} \equiv \underline{S}_{cl} \times \frac{\partial \mathcal{H}}{\partial \underline{S}_{cl}} \tag{106}$$

which is that in (47) (putting back the factor γ).

What of fluctuations around this classical limit? We see from what we're just done that we've achieved the classical limit, not by letting $\hbar \to 0$, but instead by letting $S \to \infty$. Thus we must handle fluctuations in a "$1/S$-expansion" about

the classical path S_{cl}; we do this later in this chapter, when we look at $1/N$ expansions.

Before leaving the single spin for the moment, let's look at 2 other striking features of its behaviour. First, let's consider the precessional motion noted in the context of the simple Zeeman Hamiltonian (98). Suppose we think of this problem in the path integral language we have developed here. Then we see that we are dealing with a particle of charge S moving on the surface of a sphere which has equipotentials like those shown in the figure; and in the field of a unit central monopole. Now notice that the motion in the presence of the monopole is completely different from what it would be without it, whether we deal with the limit $S \gg 1$ or not. Without the monopole, the particle would simply move accross the equipotentials (and settle at lowest point, in the direction of $\underline{B}_0$, provided there was a dissipative mechanism); with no dissipation it would simply oscillate around this direction. However the monopole completely changes - now it moves along the equipotentials, at least in the classical limit, as (106) shows. For small S, we shall see that paths going all over the Bloch sphere are allowed - but even a spin-$\frac{1}{2}$ system shows perfect precessional motion, as NMR shows. We will come back to this point.



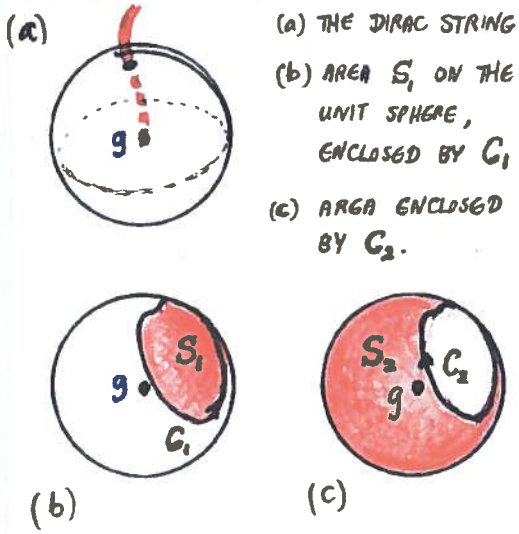
Equipotentials on the Bloch sphere for a spin in a field $\underline{B}_0$, constant in time.

The second striking feature comes when we compare integer and $\frac{1}{2}$-integer spins. The key point - and it is elementary - is that the solid angle on the Bloch sphere is only defined modulo 4π. This is shown in the figure - an area is ambiguous, and can have either sign. Thus we have invariance of all physical quantities under the transformation

$$\omega_c \rightarrow \omega_c + 4\pi n\pi \qquad (107)$$

where $n = \pm 1, \pm 2$, etc. But it then follows that

$$e^{4\pi i n S} = 1 \quad \Longrightarrow \quad S = \begin{matrix} n \\ n+\frac{1}{2} \end{matrix} \qquad (108)$$



ie., that the only allowed values of S are integer or half-integer. This argument does not of course tell us the connection between spin and statistics.

LATTICE OF SPINS:

Let's now consider a lattice of identical spins, with spin S each. In the real world there are 1-d, 2-d, and 3-d lattices of this kind, as well as interesting hybrids. Study of these is a huge field, partly because the variety of Hamiltonians is very large. We will only do a few things here, which are of

interest in a general field-theoretical context. We first take the lattice Hamiltonian and look at it in a long-wavelength continuum limit - this is done heuristically rather than rigourously. It is then shown how one can very easily derive simple field theories for both ferromagnetically & antiferromagnetically coupled spin systems, in a path integral framework. The models turn out to be well known in relativistic QFT, and have played a key role in strong interaction physics over the years.

To begin with let's consider a very simple isotropic exchange Hamiltonian, & how to describe the classical limit and the small fluctuations around it. This will done first of all in a condensed matter style, and then redone in a more field-theoretical style. I will keep the details reasonably brief.

We consider an isotropic exchange Hamiltonian, for a simple lattice of spins, which we will take to be a cubic lattice (in 3d) or a square lattice (in 2d). The interactions are between nearest neighbours; the Hamiltonian is

$$\mathcal{H} = -J \sum_{\langle ij \rangle} \underline{S}_i \cdot \underline{S}_j \tag{109}$$

where $\sum_{\langle ij \rangle}$ signifies a sum over nearest-neighbour sites, and we have 2 cases, viz., (i) the ferromagnetically-coupled (FM) case $J > 0$, and (ii) the antiferromagnetic (AFM) case $J < 0$. The action for the system is

$$S[\{\underline{n}_j\}] = \hbar S \sum_j \omega_{21}^j [\underline{n}_j] + JS^2 \int_1^2 dt \sum_{\langle ij \rangle} \underline{n}_i(t) \cdot \underline{n}_j(t) \tag{110}$$

where the vector potential $\underline{A}_j = \underline{A}$ is assumed to be the same for each site, and so

$$\omega_{21}^j [\underline{n}_j] = \int_1^2 d\underline{n}_j(t) \cdot \underline{A} \tag{111}$$

for the topological phase associated with the j-th site. The initial and final states are different for each spin, in general; we have therefore an amplitude $G(2,1)$ of form.

$$G(2,1) = \prod_{j=1}^{N} G_j(2,1) = \prod_{j=1}^{N} \int_{\underline{n}_j(t_1)}^{\underline{n}_j(t_2)} \mathcal{D}\underline{n}_j(t)\, e^{\frac{i}{\hbar} S[\{\underline{n}_j\}]} \tag{112}$$

and the vacuum-to-vacuum amplitude is

$$G(0,0) = \prod_{j=1}^{N} (2S+1)^{2N} \int \frac{d\underline{n}_j(t_2)}{4\pi} \int \frac{d\underline{n}_j(t_1)}{4\pi} \langle 0 | \underline{\mathbf{n}}(t_2) \rangle \langle \underline{\mathbf{n}}(t_1) | 0 \rangle \int_{\underline{n}_j(t_1)}^{\underline{n}_j(t_2)} \mathcal{D}\underline{n}_j(t)\, e^{\frac{i}{\hbar} S} \tag{113}$$

where $|0\rangle_N$ is the ground state of the entire system, and $|\underline{\mathbf{n}}(t)\rangle \equiv \prod_{j=1}^{N} |\underline{n}_j(t)\rangle$.

For the 2 states (FM and AFM) that we will look at, it is a fairly simple matter to guess what the relevant classical states must be; and it is also pretty simple to calculate the quadratic fluctuations we around these

classical states. We will do this in two ways:

(a) Using Holstein-Primakoff operators. These are described in Appendix B.5; for our purposes we simply need to know that once we have defined the relevant classical state, we defined the spin fluctuations around it using a set of operators $\{b_j, b_j^+\}$, which are bosons, and defined at each site. In the simplest FM case, where all spins are parallel along the z-axis in the classical ground state, we have $\underline{S}_j^{(cl)} = S\hat{\underline{z}}, \forall j$, and the fluctuations are

$$\left.\begin{aligned} S_j^z &= S - b_j^+ b_j \\ S_j^+ &= (2S)^{1/2}\left[1 - \frac{1}{2S} b_j^+ b_j\right]^{1/2} b_j \\ S_j^- &= (2S)^{1/2} b_j^+ \left[1 - \frac{1}{2S} b_j^+ b_j\right]^{1/2} \end{aligned}\right\} \quad (114)$$

where we notice the square-root reduction (this because $S_z^2 + S_x^2 + S_y^2 = S^2$) in $S_j^\pm$). The representation for the AFM case is a little more complicated, but basically the same - it is described below.

For a uniform lattice system it is more useful to Fourier transform these relations; then, eg.,

$$S^z(\underline{r}_j) = S - \frac{1}{V}\sum_{\underline{k}\,\underline{q}} e^{-i\underline{q}\cdot\underline{r}_j} b_{\underline{k}+\underline{q}}^+ b_{\underline{k}} \quad (115)$$

and so on; here V is the system volume. One can then rewrite that Hamiltonian in this language, and since we already have the classical state, it is merely a question of diagonalizing the Hamiltonian to find the fluctuation spectrum — we do this below.

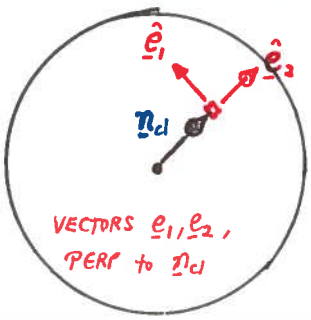
(b) Using Path Integrals. In this case we must start again from (110), and divide the vectors $\{\underline{n}_j(t)\}$ into their classical part and the fluctuation part — we will therefore write

$$\underline{n}_j(t) = \underline{n}_j^0(t)\left[1 - \underline{\phi}_j^2\right]^{1/2} + \sum_{\alpha=1,2} \hat{\underline{e}}_{j\alpha}(t)\phi_j^\alpha(t) \quad (116)$$

or, in a continuum approximation

$$\underline{n}(x) = \underline{n}_0(x)\left[1 - \underline{\phi}^2(x)\right]^{1/2} + \hat{\underline{e}}(x)\cdot\underline{\phi}(x) \quad (117)$$

where the pair of vectors $\underline{e}_1(x), \underline{e}_2(x)$ is perpendicular to the local axis $\underline{n}_0(x)$ of the spin, defined by the classical theory (see figure). The latter formulation in (117) is deliberately written in relativistic notation - the field $\underline{\phi}(x) = (\phi_1(x), \phi_2(x), \phi_3(x))$ describes small fluctuations around the classical field $\underline{n}_0(x)$; here $x = (\underline{r}, t)$ is the $(D+1)$-dimensional space in which the spin system exists. We write the gradients of these fields as



$$\left.\begin{aligned} \partial_\mu \underline{n}_{cl}(x) &= A_\mu^\alpha \hat{\underline{e}}_\alpha \\ \partial_\mu \hat{\underline{e}}^\alpha(x) &= \Gamma_\mu^{\alpha\beta}\hat{\underline{e}}_\beta - A_\mu^\alpha \underline{n}_0 \end{aligned}\right\} \quad (118)$$

where we see that the gradient of $\underline{n}_{cl}(x)$ has to be perpendicular to $\underline{n}_{cl}(x)$ itself (because $|\underline{n}(x)| = 1$ everywhere).

Note that the vector field $\underline{\phi}(x)$ is just a vector analogue to the fluctuation field $\psi(x)$ we dealt with already, in discussing the fluctuations $\psi(x)$ around the classical solution $\phi_{cl}(x)$ of the scalar field system. Here we have made it a 2-d vector, moving on the Bloch sphere. But one can also consider a quite general N-dimensional vector field $\underline{n}(x)$, and have it move on an $(N-1)$-dimensional hypersphere, and consider fluctuations $\underline{\phi}(x)$ on this hypersphere. Then the field $\underline{\phi}(x)$ is an $O(N-1)$ vector field; eqtns. (116) - (118) go through as before, except that now the sum $\alpha = 1, 2, \ldots N-1$ in (116).

Consider now how we might write the path integrals for this theory. It is convenient to take a specific form for the Lagrangian, so we choose the continuum form of (110), expanded to lowest order in gradients; the simplest relativistic analogue of this reads, for the generating functional $Z[\underline{J}]$ with a current $\underline{J}(x)$ coupling to $\underline{n}(x)$, as (dropping a constant JS^2N):

$$Z[\underline{J}(x)] = \int \mathcal{D}\underline{n}(x)\, \delta(\underline{n}^2(x)-1)\, e^{\frac{i}{\hbar}\int d^4x \left(\left[\frac{1}{\hbar} S\omega[\underline{n}(x)] + \frac{JS^2}{2}(\partial_\mu n(x) \partial^\mu n(x)) \right] - \underline{J}(x)\cdot n(x) \right)} \tag{119}$$

where the δ-fn is inserted to keep $n^2(x) = 1$ everywhere. Now we can rewrite this path integral in 4 different interesting ways:

- We can deal with the δ-fn by introducing a new field $\lambda(x)$, and performing a functional Fourier transform to get

$$Z[J(x)] = \int \mathcal{D}\lambda(x) \int \mathcal{D}\underline{n}(x)\, e^{\frac{i}{\hbar}\int d^4x \left(\frac{1}{\hbar} S\omega[n(x)] + \frac{JS^2}{2}(\partial_\mu n \partial^\mu n) + \lambda(x)(n^2(x)-1) - \underline{J}(x) n(x) \right)} \tag{120}$$

so that we have a quadratic form for the Lagrangian, we a new field which is basically behaving like a Lagrange multiplier.

- We can also separate out the classical field $\underline{n}_0(x)$ from the fluctuation around it, and then integrate separately over the 2 fields. From (118) we see that we can eliminate the classical field entirely from the exchange term in the action, because $\underline{n}_0$ is eliminated in the calculation of $\partial_\mu \underline{n} \partial^\mu \underline{n}$:

$$\frac{JS^2}{2}\int d^4x\, (\partial_\mu \underline{n}\, \partial^\mu \underline{n}) = \frac{JS^2}{2}\int d^4x \left\{ \left[(1-\underline{\phi}^2)\,\delta^\beta_\alpha + \phi_\alpha \phi^\beta \right] A^\alpha_\mu A^\mu_\beta + (\partial_\mu \phi^\alpha - \Gamma^{\alpha\beta}_\mu \phi_\beta)(\partial^\mu \phi^\alpha - \Gamma^{\mu}_{\alpha\beta}\phi^\beta) \right\} \tag{121}$$

so that we can reduce the functional integral over the exchange term to one over the A^α_μ and $\Gamma^{\alpha\beta}_\mu$ fields. This is less useful if we also need to do the full integral over $\underline{n}$ in the other terms, but if we are only interested in the fluctuations, it gives a useful alternative.

- We can use the CP^N representation. Consider the set of complex variables $\underline{z}(x) = (z_1(x), z_2(x) \ldots z_N(x))$, with the constraint that

$$|\underline{z}^2| = \sum_{j=1}^{N} |z_j(x)|^2 = 1. \tag{122}$$

If we think of $\underline{z}(x)$ as a field, we can make the assumption that is invariant under local gauge transformations, ie., invariant under $\underline{z}(x) \to \underline{z}(x) e^{i\varphi(x)}$. We then make the standard manouevre, viz., write a free field Lagrangian

$$S_0 = \frac{1}{g_0^2} \int d^4x \, (\partial_\mu + i A_\mu) \underline{z}^\dagger (\partial^\mu - i A^\mu) \underline{z} \tag{123}$$

where $\underline{z}^\dagger(x)$ is the Hermitian conjugate of $\underline{z}(x)$; and the theory is invariant under the gauge transformation $A_\mu(x) \to A_\mu(x) + \partial_\mu \varphi(x)$. Note that there is no kinetic term for $A_\mu(x)$ in this action, so we can eliminate $A^\mu(x)$ in the classical limit by minimizing the action. This just gives

$$\frac{\delta}{\delta A_\mu} \left[(\partial_\mu \underline{z}^\dagger)(\partial^\mu \underline{z}) - i \left(A_\mu \underline{z}^\dagger \partial^\mu \underline{z} - A^\mu \underline{z} \partial_\mu \underline{z}^\dagger \right) + A_\mu A^\mu \right] = 0 \tag{124}$$

so that

$$A_\mu(x) = -\frac{i}{2} \left(\underline{z}^\dagger \partial_\mu \underline{z} - \underline{z} \partial_\mu \underline{z}^\dagger \right) \tag{125}$$

We now make the connection with the $O(3)$ $\underline{n}$-field; letting $N = 2$ in our CN^N model, we write

$$\left.\begin{aligned} \underline{n}(x) &= \underline{z}^\dagger \underline{\tau} \underline{z} \equiv z_\alpha^\dagger \tau^{\alpha\beta} z_\beta \\ &\equiv (z_1^\dagger, z_2^\dagger) \begin{pmatrix} \tau_{11} & \tau_{12} \\ \tau_{21} & \tau_{22} \end{pmatrix} \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} \end{aligned}\right\} \tag{126}$$

so that the z_α now are components of a spinor, with $\underline{\tau}$ the Pauli vector matrix; we then have

$$\frac{Js^2}{2} \int d^4x \, (\partial_\mu n \, \partial^\mu n) = 2Js^2 \int d^4x \left[(\partial_\mu + i A_\mu) \underline{z}^\dagger (\partial^\mu - i A^\mu) \underline{z} \right] \tag{127}$$

with A_μ again satisfying (125).

It is interesting to note what form the field "curvature" tensor $F_{\mu\nu}$ takes in this representation; we have

$$\left.\begin{aligned} F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu &= \underline{n} \cdot (\partial_\mu n \times \partial_\nu \underline{n}) \\ &\equiv \epsilon_{\alpha\beta\gamma} n^\alpha \partial_\mu n^\beta \partial_\nu n^\gamma \end{aligned}\right\} \tag{128}$$

which, as we will now show, is just a topological charge density for the

$\underline{n}$-field. We define the scalar quantity

$$Q = \frac{1}{4\pi}\int d^2x\, \Phi(x) = \frac{1}{4\pi}\int d^2x\, \epsilon^{\mu\nu} F_{\mu\nu}(x)$$
$$= \frac{1}{4\pi}\int d^2x\, \epsilon^{\mu\nu}\, \underline{n}(x)\left(\partial_\mu \underline{n}(x) \times \partial_\nu \underline{n}(x)\right) \qquad (129)$$

which is the "winding number of a sphere onto a sphere"; for any smoothly-varying $\underline{n}(x)$, it will be an integer. $\Phi(x)$ is called "Pontryagin density", and Φ the Pontryagin index, by mathematicians. We see that if there was a kinetic term for the gauge field of the $\underline{n}$-vector, it would be

$$T \sim \int d^2x\, F_{\mu\nu}(x) F^{\mu\nu}(x) = \Phi^2 \qquad (130)$$

Note that we can also write the topological or Berry phase term in a simple way in this CPN language. It is easy to see that if we write this in the form

$$S_B = i\hbar S \sum_j \omega^j[\underline{n}_j]$$
$$= i\hbar S \sum_j \int dt\, \underline{A}\cdot\frac{d\underline{n}_j}{dt} \longrightarrow i\hbar S \int dt \sum_j \underline{z}_j^\dagger \frac{d}{dt}\underline{z}_j \qquad (131)$$

Let us, for example, pick a form for $\underline{z}_j$ in CP_2 that is in the same gauge as we picked before, with the Dirac string through the north pole; then we have

$$\underline{z}_j(t) = \begin{pmatrix} z_j^+(t) \\ z_j^-(t)\end{pmatrix} = \begin{pmatrix} \cos\theta_j/2\, e^{-i\phi_j} \\ \sin\theta_j/2\end{pmatrix} \qquad (132)$$

and it is easily verified that (126) is correct.

Note, incidentally, the relationship between the complex 2-component vector $\underline{z}$ and the quantity $z = e^{i\phi}\tan\theta/2$ that appears in the definition of coherent states for spin-$\frac{1}{2}$ (cf. App. B5). We see that $z = z_-/z_+$, as we would expect since z refers to the Riemann projection of the $\underline{n}$-vector.

Let's now summarize all that we have found for the different representations of the path integral for a spin system. We suppose that we start with a Hamiltonian $\mathcal{H}(\{\underline{s}_j\})$, which could be of simple exchange type, or more complicated. We can then represent the system in the following ways.

(i) Using the Holstein-Primakoff expansion in (114).

(ii) In terms of the $\underline{z}$-fields, or the related coherent spin states, with an action given by

$$S[\{\underline{z}_j^\dagger, \underline{z}_j\}] = \sum_j \int dt \left[i\hbar S\, \underline{z}_j^\dagger \dot{\underline{z}}_j - \mathcal{H}(\{\underline{z}_j^\dagger, \underline{z}_j\})\right] \qquad (133)$$

where the Hamiltonian is written using (125) in terms of the $\{\underline{z}_j\}$.

(iii) Directly in terms of the $\underline{n}$-vector field, so that we can write the generating functional

$$\left.\begin{aligned} Z[\underline{J}] &= \int \mathcal{D}\underline{n}(x)\, \delta(|\underline{n}(x)|^2 - 1)\, e^{\frac{i}{\hbar}\int d^4x \left[\hbar S\, \omega[\underline{n}(x)] - \mathcal{H}[S\underline{n}(x)] - \underline{J}(x).\underline{n}(x)\right]} \\ &\equiv \int \mathcal{D}\lambda(x) \int \mathcal{D}\underline{n}(x)\, e^{\frac{i}{\hbar}\int d^4x \left[\hbar S\, \omega[\underline{n}] - \mathcal{H}[S\underline{n}] + \lambda(x)(|\underline{n}(x)|^2 - 1) - \underline{J}(x).\underline{n}(x)\right]} \end{aligned}\right\} \quad (134)$$

where we have coupled the spin to an external field $\underline{J}(x)$. This is the most obvious way to do path integrals for the spin field $\underline{n}(x)$. But we can also write things in terms of magnon "Holstein-Primakoff" operators, or in terms of the $\underline{z}(x)$ fields, and this gives real flexibility. In order to understand all of this better, we need now to look at some examples.

(ii) QUANTUM FERROMAGNET :

This is the simplest example we can consider. Let us however look at a slightly more general example than the simple isotropic model of (109); we consider instead the Hamiltonian

$$\left.\begin{aligned} \mathcal{H} &= -\gamma \sum_j H_0 S_j^z - \sum_{\langle ij\rangle} \left[J_{ij}^{\parallel} S_i^z S_j^z + J_{ij}^{\perp}(S_i^x S_j^x + S_i^y S_j^y)\right] \\ &\equiv -\gamma \sum_j H_0 S_j^z - \sum_{\langle ij\rangle} \left[J_{ij}^{\parallel} S_i^z S_j^z + \tfrac{1}{2} J_{ij}^{\perp}(S_i^+ S_j^- + S_i^- S_j^+)\right] \end{aligned}\right\} \quad (135)$$

ie., we have an external field applied along the z-axis, and we have 2 different exchange constants, retaining the longitudinal and transverse terms (ie., we have "exchange anisotropy"). Note that this far from being the most general ferromagnetically-coupled Hamiltonian that one could consider *

In what follows I will not treat this problem using path integral methods, but instead use the old-style method of Holstein-Primakoff operators, which will give us a very straightforward path into the problem.** We begin by asking what the ground state of the system ought to be; this will also be our classical starting point.

Note first that in the absence of the exchange terms, the ground state of the system is obvious, no matter what is S (the spin at each site); we just have the state

$$|0\rangle = \prod_{j=1}^{N} |S, m_j = S\rangle \equiv |\uparrow\uparrow\uparrow \ldots \uparrow\rangle \quad (136)$$

in which all spins are pointing up along the z-axis (for $H_0 > 0$). Now let's ask

---

* Important terms missing from (129) include:

(i) Single-ion terms, of form $\sum_j K_j^{\alpha\beta\gamma\ldots} S_j^\alpha S_j^\beta S_j^\gamma \ldots$ (a simple one is $\sum_j K_j^{\parallel} (S_j^z)^2$).

(ii) More complex "multi-site" exchange terms, of form $\sum_{ijkl\ldots} J_{ijkl\ldots}^{\alpha\beta\gamma\ldots} S_i^\alpha S_j^\beta S_k^\gamma \ldots$

(iii) Long-range inter-spin dipolar interactions, which don't even conserve total spin.

** Done by Holstein & Primakoff in 1940; one of the first uses of QFT in condensed matter.

what happens when we add the exchange term, noting that for FM ordering we require that $J_{ij}^{\parallel}$ and $J_{ij}^{\perp} > 0$. From ordinary QM it is clear that this does not disturb the state $|0\rangle$, because

$$\sum_j \left[S_i^z ,\ S_i^x S_j^x + S_i^y S_j^y \right] \quad = \quad 0 \tag{137}$$

We therefore establish that $|0\rangle$ is still an eigenstate of the full $\mathcal{H}$. Whether it is still the ground state needs a little more investigation. Notice first that we can think of the Hamiltonian, apart from the external field Zeeman term, as a set of pairwise interactions, so it suffices to understand what happens to a single pair. Second, we see that we can also rewrite (135) as

$$\mathcal{H} = -\gamma \sum_j H_0 S_j^z - \sum_{\langle ij \rangle} \left[J_{\perp}\, \underline{S}_i \cdot \underline{S}_j + (J_{\parallel} - J_{\perp})\, S_i^z S_j^z \right] \tag{138}$$

from which we see that if $(J_{ij}^{\parallel} - J_{ij}^{\perp}) > 0$, the last term merely increases the stability of the state $|0\rangle$, and so all we have to worry about is the isotropic term, which we can rewrite as

$$\begin{aligned} \mathcal{H}_{\text{isotr.}} &= -\sum_{\langle ij \rangle} \frac{J_{\perp}}{2} \left[(\underline{S}_i + \underline{S}_j)^2 - \underline{S}_i^2 - \underline{S}_j^2 \right] \\ &= -\sum_{\langle ij \rangle} \frac{1}{2} J^{\perp} \left[2S(S+1) - (\underline{S}_i + \underline{S}_j)^2 \right] \end{aligned} \tag{139}$$

and the ground state of this is obviously $|0\rangle$ as well, since the energy is minimized by making the spins parallel. Thus the only real difficulty arises when $(J_{\parallel} - J_{\perp}) < 0$; and in fact one can destabilize the simple ground state to other more complicated states under these circumstances. Thus we have the following possibilities.

(a) When $H_0 = 0$, and $J_{\parallel} = J_{\perp} = J$, so that

$$\mathcal{H} = -\sum_{\langle ij \rangle} J\, \underline{S}_i \cdot \underline{S}_j \tag{140}$$

the ground state has all spins parallel; however, this state is actually $(2NS+1)$-fold degenerate (where N is the number of spins), because the total spin $\mathbf{S} = \sum_j \underline{S}_j$ can point in any direction; the system is invariant under rotations in spin space, so that the correct ground state is a symmetric combination of all of these, so that in a coherent state representation we just have

$$|0\rangle_{\text{symm}} = \int \frac{d\Omega_N}{4\pi} |\underline{N}\rangle \quad \equiv \quad \prod_{j=1}^{N} \int \frac{d\eta_j}{4\pi} |\underline{\eta}_j\rangle\, \delta(\underline{N} - \underline{\eta}_j) \tag{141}$$

where $|\underline{N}\rangle$ is the state defined after (113), and $\underline{N}$ is the direction of $|\underline{N}\rangle$. Basically what we have here is a coherent superposition of magnetizations pointing in all directions. — very far from being a

classical state! However...

(b) When $H_0 \neq 0$, we immediately change this situation. Thus for the isotropic exchange Hamiltonian with applied field, ie., for

$$\mathcal{H} = -\gamma \sum_j H_0 S_j^z - \sum_{<ij>} J \underline{S}_i \cdot \underline{S}_j \qquad (142)$$

the ground state $|0\rangle$ in (130) quickly drops below all the other states involved in the superposition. To see what the energy gap is to the first excited state, we might think of trying the state

$$|1_k\rangle = \left(\prod_{i \neq k}^{N} |\uparrow_i\rangle \right) |\downarrow_k\rangle \qquad (143)$$

which has an energy $2\gamma S$ higher than $|0\rangle$; but it is physically obvious (and will become quantitatively clear below) that we can get an excited state with much lower energy by delocalizing the flipped spin in (143). To properly discuss the definition of the classical state and the excited states for this system, we need to look at fluctuations, which we do below.

(c) Finally, for the full Hamiltonian in (135), $|0\rangle$ will remain the ground state provided $(J_{\parallel} - J_{\perp}) > 0$; otherwise things get more complex.

HOLSTEIN-PRIMAKOFF TREATMENT: Let's look first at the isotropic system (109), with an applied field - the full case in (135) involves exactly the same manoeuvres, but is more messy. By substituting the Holstein-Primakoff operators into (109), and making the expansion of the square roots, so that, eg.,

$$\begin{aligned} S_j^+ &= (2S)^{1/2} \left[1 - \frac{1}{2S} b_j^\dagger b_j \right]^{1/2} b_j \\ &= (2S)^{1/2} \left[b_j - \frac{1}{4S} b_j^\dagger b_j b_j + O(1/S^2) \right] \\ &\equiv \left(\frac{2S}{N} \right)^{1/2} \left[\sum_q e^{-i\underline{q}\cdot\underline{j}} b_q - \frac{1}{4NS} \sum_{q_1 q_2 q_3} \delta e^{i(q_1 - q_2 - q_3)} b_{q_1}^\dagger b_{q_2} b_{q_3} + O(1/S^2) \right] \end{aligned} \qquad (144)$$

we can write the Hamiltonian in lattice space as

$$\mathcal{H} = \gamma H_0 \sum_j b_j^\dagger b_j - 2S^2 \sum_{<ij>} J_{ij} \qquad (145)$$

where we subtract off the constant ground state energy, given by

$$E_0 = -\gamma N S H_0 - \frac{1}{S} \sum_{<ij>} S^2 J_{ij} \left(b_i^\dagger b_j - b_j^\dagger b_i \right) + \cdots \qquad (146)$$

where the higher terms are higher order in $1/S$. We can also write this in terms

of the Fourier transformed operators $b_q, b_q^\dagger$; defining the Fourier transform of the exchange interaction as

$$J_q = \sum_{j \neq i} J_{ij} e^{i q \cdot (r_i - r_j)} \xrightarrow[n.n.]{} J z \gamma_q \tag{147}$$

where

$$\gamma_q = \frac{1}{z} \sum_{\langle nn \rangle} e^{i q \cdot r_{nn}} \tag{148}$$

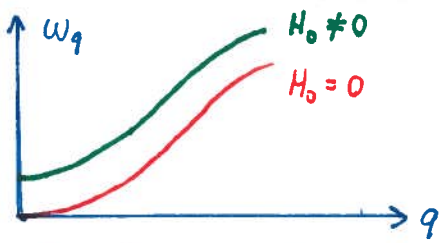
is the "form factor" and z, the coordination number, is the number of nearest neighbor around any given lattice site; we can now write (145) as

$$\mathcal{H} = \gamma H_0 \sum_q n_q + z S \sum_q J (1 - \gamma_q) n_q \tag{149}$$

where to simplify things we've assumed inversion symmetry in the lattice, so that $J_q = J_{-q}$, and where $n_q = b_q^\dagger b_q$. Thus we can write $\mathcal{H}$ in the form

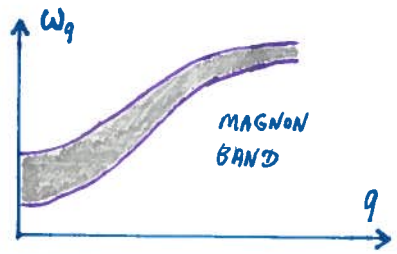
$$\mathcal{H} = \sum_q \omega_q^0 b_q^\dagger b_q \tag{150}$$

where the magnon dispersion relation goes like



ABOVE: MAGNON DISPERSION IN ISOTROPIC SYSTEM

BELOW: FOR ANISOTROPIC SYSTEM



$$\omega_q^0 = z S J (1 - \gamma_q) + \gamma H_0 \tag{151}$$

which for small q must be $\propto |q|^2$; in fact we have (with $c \sim O(1)$):

$$\left.\begin{aligned} \omega_q^0 &\sim \gamma H_0 + \tfrac{1}{2} \sum_{\langle nn \rangle} (q \cdot r_{nn})^2 + O(q r_{nn})^4 \\ &\sim \gamma H_0 + c J S (q a_0)^2 + O(q a_0)^4 \end{aligned}\right\} \tag{152}$$

so that we always have a mass gap induced by the applied, which goes to zero with the field, and a (non-relativistic) form for the magnon dispersion above this gap.

The detailed form for γ_q depends on the crystal structure. Here are a few important examples:

1 dimension ($z = 2$): Then we have: $\gamma_q = \cos q a_0$ (153)

2 dimensions: ($z = 4$): For a square lattice $\gamma_q = \frac{1}{2}(\cos a_0 q_x + \cos a_0 q_y)$ (154)

3 dimensions: The most common are:

- Simple cubic ($z = 6$): $\gamma_q = \frac{1}{3}(\cos q_x a_0 + \cos q_y a_0 + \cos q_z a_0)$
- B.C.C. ($z = 8$): $\gamma_q = \frac{1}{8}\left(\cos \frac{q_x a_0}{2} \cos \frac{q_y a_0}{2} \cos \frac{q_z a_0}{2}\right)$
- F.C.C. ($z = 12$): $\gamma_q = \frac{1}{3}\left[\cos \frac{q_x a_0}{2} \cos \frac{q_y a_0}{2} + \cos \frac{q_x a_0}{2} \cos \frac{q_z a_0}{2} + \cos \frac{q_y a_0}{2} \cos \frac{q_z a_0}{2}\right]$

(155)

These results illustrate one key general result, and also raise a key question. The key result is that as $\gamma H_0 \to 0$, we end up with a set of "Goldstone modes", viz., modes whose frequency $\to 0$ as $q \to 0$. We will not discuss this point in great detail here - it is analyzed later in the context of broken symmetries and order parameters. However, very roughly speaking, we can say that in a situation where a symmetry of the underlying Hamiltonian is "spontaneously broken" in the ground state, then there will be a branch of excitations which has frequency $\to 0$ as $q \to 0$. This is a delicate point, because we saw that when $H_0 = 0$ rigorously, the ground state is actually a superposition of form (141), ie, there is no broken symmetry. The way this is conventionally formulated is using a specific order of limits; we say that

$$\lim_{H_0 \to 0} \lim_{N \to \infty} |0\rangle = |\uparrow\uparrow\uparrow \cdots \uparrow\rangle \qquad (156)$$

to replace (141), where we take the thermodynamic limit $N \to \infty$ before taking the applied field to zero. In the opposite case, where we first take H_0 to zero, we have

$$\lim_{N \to \infty} \lim_{H_0 \to 0} |0\rangle = |0\rangle_{symm} = \int \frac{d\Omega_{\hat{N}}}{4\pi} |\hat{N}\rangle \qquad (157)$$

ie., the totally symmetric state in (141).

The question raised by these results is simply this - how trustworthy are they? We notice that they are obtained to leading order in a $1/S$ expansion, so we might expect them to be more accurate for large S, and this indeed so. But things are actually more subtle than this. In fact a more careful statement of the Goldstone result is

> "If an underlying symmetry is spontaneously broken, and the interparticle/interspin couplings are sufficiently short-ranged, then a Goldstone mode ($\omega_q \to 0$ as $q \to 0$) will appear". (158)

However this does not tell us when the symmetry will be broken. We will return again to this point later, once we have looked at the AFM system; but suffice it to say that the result depends both on the form of the interactions and on the dimensionality of the system. To give a foretaste of what we will see, and to clarify a little the meaning of (158), let's extend very slightly the treatment being given here, in 2 ways:

(a) HIGHER MAGNON TERMS: The effective Hamiltonians in (149) and (150) are produced by the lowest-order expansion in $1/S$; they give a set of non-interacting magnons, ie, a free field theory with small oscillations around the ground state, in which these oscillations correspond to a quadratic form in the $b_q, b_q^\dagger$ and are therefore uncoupled. From this it is completely obvious how to write the action, propagators, and diagram rules for this free field theory, with and without the field H_0.

However as soon we go to higher order in $1/S$, interactions between the magnons are generated. The form of these terms and of their interaction kernels is messy, & depends in a complicated way on the interactions in the original Hamiltonian. Because of this complexity we only make a few brief remarks:

* If we simply take the isotropic exchange FM, then further expansion in powers of $1/S$ produces interactions between even numbers of magnons only. Thus we get a series sum of form

$$\mathcal{H} = \sum_{q} \hbar\omega_q^0 b_q^\dagger b_q + \sum_{n=2}^{\infty} \sum_{q_1 \cdots q_{2n}} \Gamma^{(2n)}_{q_1 \cdots q_{2n}} b_{q_1}^\dagger b_{q_2}^\dagger \cdots b_{q_n}^\dagger b_{q_{n+1}} \cdots b_{q_{2n}} \delta\left(\sum_{\alpha} q_\alpha\right) \tag{159}$$

where the vertices $\Gamma^{(2n)}_{q_1 \cdots q_{2n}}$ have to be calculated for a given lattice symmetry and form for J_{ij}. The results can be complicated; thus, for a simple cubic lattice and nearest-neighbor J, one finds

$$\left.\begin{aligned} \Gamma^{(4)}_{q_1 \cdots q_4} &= \frac{Jz}{2S}\left[\left(\gamma_{q_1} + \gamma_{q_4} - 2\gamma_{q_2 - q_4}\right) + \frac{1}{8}\left(\gamma_{q_1} + q_{q_4}\right)\right] \\ \Gamma^{(6)}_{q_1 \cdots q_6} &= \frac{Jz}{16 S^2}\left[\gamma_{q_1} + \gamma_{q_6} - 2\gamma_{q_3 - q_5 - q_6}\right] \end{aligned}\right\} \tag{160}$$

and so on. Thus, diagrammatically, we have generated a theory with an infinite set of interactions, reminiscent of our discussion of phonons in a solid.

* As soon as rotational symmetry about the magnetization axis is broken, vertices with odd numbers of magnons also enter the effective Hamiltonian. Thus, a much more general Hamiltonian for the system involves external fields, exchange & dipolar interactions, and both exchange & single-ion anisotropy terms; a simple example would be one with 2-site exchange and quadratic single-ion anisotropy:

$$-\mathcal{H} = -\sum_j \gamma_j^{\alpha\beta} H_0^\alpha S_j^\beta + \sum_j K_j^{\alpha\beta} S_j^\alpha S_j^\beta - \sum_{ij} J_{ij}^{\alpha\beta} S_i^\alpha S_j^\beta + \sum_{ij} V_{ij}^{\alpha\beta} S_i^\alpha S_j^\beta \tag{161}$$

where $\gamma_j^{\alpha\beta} = \mu_B g^{\alpha\beta}$ is an anisotropic g-factor for the spins, $K_j^{\alpha\beta}$ is the single-ion anisotropy tensor, and the dipolar interaction $V_{ij}^{\alpha\beta}$ takes the form

$$\left.\begin{aligned} V_{ij}^{\alpha\beta} &= \frac{\mu_B^2}{r_{ij}^3}\left[\delta_{\mu\nu} - 3\frac{r_{ij}^\mu r_{ij}^\nu}{r_{ij}^2}\right] g_i^{\mu\alpha} g_j^{\nu\beta} \\ &\xrightarrow[\text{limit}]{\text{isotropic}} \left[\delta_{\alpha\beta} - 3\frac{r_{ij}^\alpha r_{ij}^\beta}{r_{ij}^2}\right] V_0(r_{ij}) \end{aligned}\right\} \tag{162}$$

where in the isotropic limit, $\mu_B g^{\alpha\beta} \to \mu_B g \delta^{\alpha\beta} = \gamma$; then in this

limit, the total dipolar interaction takes the familiar form:

$$V_{ij}^{\alpha\beta} S_i^\alpha S_j^\beta \xrightarrow[\text{isotropic}]{} V_0(r_{ij}) \left[\underline{S}_i \cdot \underline{S}_j - 3 \frac{(\underline{r}_{ij} \cdot \underline{S}_i)(\underline{r}_{ij} \cdot \underline{S}_j)}{r_{ij}^2} \right] \tag{163}$$

with $\underline{r}_{ij} = \underline{r}_i - \underline{r}_j$, and the interaction strength: $V_0(\underline{r}) = \gamma^2 / |\underline{r}|^3$ (164)

If we calculate a magnon expansion with a Hamiltonian like (155) we get very messy results, which can be found in specialized papers and reviews. They are of no particular interest to us here, except to note that in general they do not conserve magnon number. Thus, e.g., suppose we consider an isotropic exchange interaction and an isotropic dipolar interaction, but with no single-ion terms and no external field. Then the result for the new effective Hamiltonian becomes

$$\left.\begin{aligned} \mathcal{H} = \sum_q \left[\omega_q^0 b_q^+ b_q + (U_q b_q b_{-q} + \text{H.c.}) \right] \\ + \sum_{q_1 q_2 q_3}{}' \left(\Gamma^D_{q_1 q_2 q_3} b_{q_1} b_{q_2} b_{q_3}^+ + \text{H.c.} \right) \delta(q_1 + q_2 - q_3) + \cdots \end{aligned}\right\} \tag{165}$$

where the coefficients U_q, $\Gamma^D_{q_1 q_2 q_3}$, etc., are complex and depend in a complicated way on the sample shape, the lattice structure, and the anisotropic g-factor. The non-conservation of magnon number reflects the fact that the dipolar interaction itself does not even conserve total spin. One can of course rediagonalize the "bare" magnon term, by a rotation in the space of operators b_q and b_{-q}^+; but the result is a set of magnons which are superpositions of the two (and thus, a superposition of $\hat{S}_q^+$ and $\hat{S}_{-q}^-$), rather than just $\hat{S}_q^+$ and $\hat{S}_q^-$ separately.*

(b) DEMAGNETIZATION FIELDS & BROKEN SYMMETRY: It is intuitively and physically obvious that when the system of spins polarized, there is an extra magnetic field in the system which is generated by the spins themselves. This term is absent from the Hamiltonian in explicit form, and absent from the effective Hamiltonians given up to (159), because its microscopic cause is to be found in the dipolar interaction. Thus, the extra field – which for historical reasons is called the demagnetization field – arises self consistently from the magnetization itself. In the continuum limit we can rewrite the dipolar term as

$$\left.\begin{aligned} \mathcal{H}_{dip} &= \sum_{ij} V_{ij}^{\alpha\beta} S_i^\alpha S_j^\beta \\ &= -\mu_0 \int d^3r \, \underline{H}_{DM}(\underline{r}) \cdot \underline{M}(\underline{r}) \\ &= \mu_0 \int d^3r \int d^3r' \, \nabla_{\underline{r}'} \cdot \underline{M}(\underline{r}') \frac{1}{|\underline{r} - \underline{r}'|} \nabla_{\underline{r}} \cdot \underline{M}(\underline{r}) \end{aligned}\right\} \tag{166}$$

* This transformation, introduced by Holstein & Primakoff in 1940, was re-used in 1947 in modified form by Bogolubov to superpose particle and anti-particle boson operators, and by BCS to superpose particle & hole fermion operators. It is now known as a Bogolubov transformation.

and we see that demagnetization field $\underline{H}_{DM}(\underline{r})$ has a very interesting non-local form, depending on the integral of $M(r)$ throughout the sample:

$$\underline{H}_{DM}(\underline{r}) = -\int d^3r' \, \underline{M}(\underline{r}') \cdot \nabla_{\underline{r}'} \frac{1}{|\underline{r}-\underline{r}'|} \nabla_{\underline{r}} \qquad (167)$$

The non-local character arises from the long-range nature of the dipolar interaction, and the form of (166), with the dependence on $(\underline{r}-\underline{r}')$, arises directly from the specific form of (163) and (164).

We can now look at this result in the context of the statement about Goldstone bosons given in (158). The basic point is this — if, in the absence of dipolar interactions, we reduce $\underline{H}_0 \to 0$, then depending on the order of limits in (156) and (157), we may or may not get ordering of the spins into a macroscopic magnetization, but this magnetization will not feed back on the system, and we will have a set of gapless Goldstone bosons, in the form of magnons with dispersion $\omega_q^0 \propto |\underline{q}|^2$. However, regardless of how we let $\underline{H}_0 \to 0$, once the dipolar interaction is introduced, the long-range nature of this will lead to a feedback effect in which a net field is produced by the spins, in a self-consistent way in which spins act to line each other up. We are all familiar with this — it is the phenomenon which leads to ordinary magnets.

However we now notice that this result has had a profound effect. First, it has destroyed the gapless Goldstone modes — the demagnetization field now induces a gap in the magnon spectrum, even when $\underline{H}_0 = 0$. This is quite consistent with (158), with its proviso that forces need to be short-ranged.

Second, we notice that what this field $\underline{H}_{DM}(\underline{r})$ will be cannot be specified unless one specifies what is the shape of the system. This crucial point is often ignored or misunderstood in the literature — it means that the system has no well-defined thermodynamic limit! If we let $N \to \infty$, the physical properties still depend on the shape of the system, via (166). Thus you should beware of much of what passes for proof in the statistical mechanics or QFT literature, which often ignores boundary conditions at infinity — given that gauge fields are intrinsically long-ranged, and Nature is described in a fairly deep way by gauge field theories, we are rarely entitled to be so cavalier about what happens at boundaries.

Notice, however, that the existence of this additional field does have a very useful effect. It provides a mechanism to stabilize the classical solution to our QFT; the gap in the spectrum makes it much more difficult for quantum fluctuations to alter the ground state. To see this in detail, and to prove it, is not all that simple, but we shall see that at least in some cases — including this one, for dimensionality $d \geq 2$, in the absence of other destabilizing influences — the idea works rather well, and the classical field will simply be the local magnetization $\underline{M}(\underline{r},t)$.

Note also that we have ignored one other key point. This is that one can develop non-trivial inhomogeneous configurations in $M(\underline{r},t)$, which are

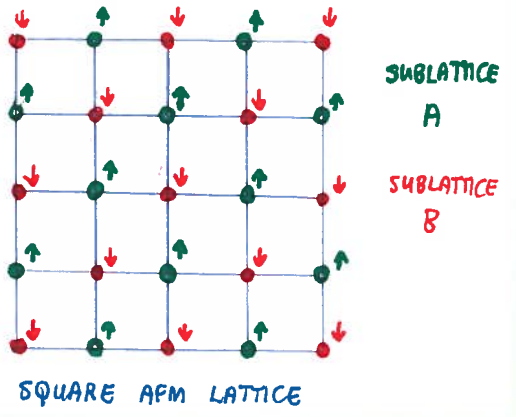
topologically stable. For a FM system, we have already seen how this can work in 1 dimension; the 1-d Sine-Gordon system is a very simple example of a system in which gapped small oscillations exist in an ordered system, which can also possess topological solitons. If we go to 2 spatial dimensions, the system can support both line domain walls and vortex solutions; and in 3 dimensions we have sheetlike domain walls, vortex lines, and "Bloch point" or "hedgehog" monopole soliton solutions.

The role and behaviour of soliton and instanton solutions in a QFT is studied in more detail later in this section.

(iii) QUANTUM ANTIFERROMAGNET :

The quantum FM is an example of a system where the classical solution to the QFT works very well - indeed, we were able to perform a fluctuation expansion around this theory which works very well for dimensionality ≥ 2. The main reason this worked so well was that the classical field stabilized itself, using the demagnetization field it generated.

The antiferromagnetically-ordered (AFM) state is quite different, in a number of different ways, which will lead us into further insights into general features of QFT. The classically ordered state was first discussed by Néel in 1932.



Néel's basic idea can be seen in the figure, which shows a 2-d square AFM array of spins, with nearest neighbours linked by vertical or horizontal lattice vectors. We see that if we divide the system into 2 sublattices, then any spin on sublattice A only interacts with spins on sublattice B, and vice-versa. With AFM interactions, ie., with $J < 0$, then it is clearly classically advantageous to pick opposite orientations for the spins on the 2 sublattices, to minimize the energy; indeed, for the simple system shown in the figure, we have a classical ground state energy $E_0 = \frac{1}{2} z J N S^2$, where $z = 4$ is the number of nearest neighbours, and we assume $J_{ij} = J$ for nearest neighbours, and zero otherwise.

However there are some obvious problems with this simple picture. The first is that, unlike the FM case, the dynamics of a spin in such an environment are not consistent with the ground state we have chosen. We can see this by just by considering a single pair of spins with AFM coupling; then the Hamiltonian is

$$\mathcal{H}_{12} = -J \underline{S}_1 \cdot \underline{S}_2 \qquad (J<0) \tag{168}$$

which has the ground state

$$|0\rangle = \frac{1}{\sqrt{2}} \left[|S,-S\rangle - |-S,S\rangle \right]_s \longrightarrow \frac{1}{\sqrt{2}} \left[|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle \right] \quad \text{spin-}\tfrac{1}{2} \tag{169}$$

where the notation $|S,S'\rangle_{\hat{z}} \equiv |S,s\rangle|S,s'\rangle$, ie., we only write the projection along $\hat{z}$ inside the ket. The energy of the ground state is

$$E^{\circ}_{12} = JS(S+1) \qquad (J<0) \tag{170}$$

Now we see that the classical state $|S,-S\rangle$ is neither the ground state nor an eigenstate; the classical energy is $JS^2 > E^{\circ}_{12}$, and

$$\mathcal{H}_{12}\,|S,-S\rangle_{\hat{z}} = J\,|-S,S\rangle_{\hat{z}} \tag{171}$$

or, for spin ½,

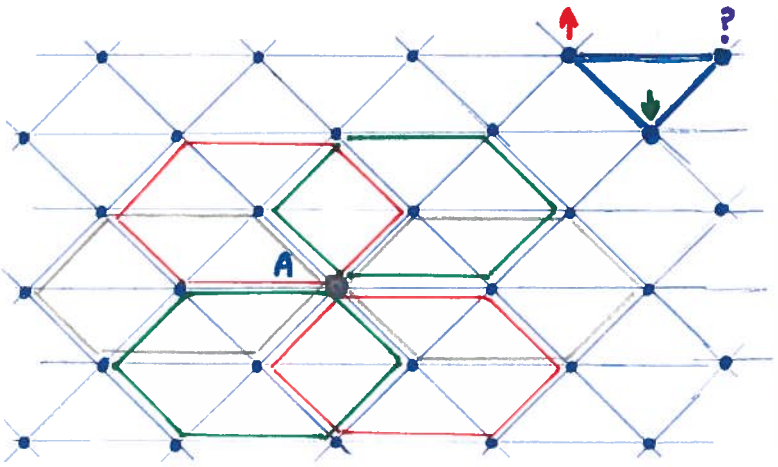
$$\mathcal{H}_{12}\,|\uparrow\downarrow\rangle = J\,|\downarrow\uparrow\rangle \tag{172}$$

These remarks simply underline the role of quantum fluctuations; the ground state is a superposition of the two possible classical states $|S,-S\rangle$ and $|-S,S\rangle$, and indeed simply resonates/oscillates between them with frequency $J/\hbar$. To put it another way, we are forming a coherent superposition of oppositely oriented spin states in which, however, the direction of the spins is entirely undetermined. Thus, in the absence of some other field which can stabilize the orientation, the classical state is a bad approximation.

At first glance these remarks seem to parallel those made about the FM state (cf. eqtn. (139) and discussion thereafter). However, the AFM situation is very different, because in the FM case, the triplet state $|\uparrow\uparrow\rangle$ is an eigenstate as well as being the classical solution. Thus, the interactions are not destabilizing the FM structure at the microscopic level, as in the AFM case.

For this reason the proposal of Néel in 1932 was initially treated rather sceptically by theorists, particularly those of the London school. We will come back to this point, which can be put in the form of a question, viz., why should the classical AFM state be stable at all? Yet, remarkably, it is – in both 3-d and 2-d systems – classical analysis works quite well, as shown by experiments. However the story turns out to have an interesting subtle twist.

The second obvious problem with the classical argument given above is that, unlike the case of the FM, it relies heavily on the lattice structure – for many lattices one cannot find an obvious classical AFM state. Consider, eg., the hexagonal lattice shown in the figure. To figure out a sublattice structure is not so obvious (and would depend on the precise form of the J_{ij}); we note that any particular lattice site partakes of 6 different hexagonal plaquettes. Suppose that $J_{ij} = J$ for nearest neighbors, with $J<0$. Consider now the energy minimization for a single triangle



HEXAGONAL LATTICE: ANY LATTICE POINT A PARTAKES OF 6 DIFFERENT HEXAGONS. EACH TRIANGLE SHOWS MAGNETIC FRUSTRATION FOR NEAREST NEIGHBOUR AFM INTERACTIONS.

of 3 spins. It is clear that we cannot minimize the energy by choosing oppositely oriented spins (if one is up, the other down, then how do we choose the 3rd one?), and in fact the classical energy is minimized by having the spins oriented at mutual angles of 120°. Again, however, the ground state of this system is not this classical state, but a superposition of them. Moreover, as one easily sees, any attempt to fit these "triangle" states together into an ordered lattice with spins oriented at 120° between nearest neighbours, is again frustrated by the lattice structure. This "frustration" effect is of great importance in magnetism — it prevents us from finding an ordered classical state.

Investigation of these issues over the last 60 yrs has led to some very interesting results. In what follows we will look at just 2 of them in detail, viz., (i) the classical spin wave theory for ordered AFM systems, and (ii) the non-linear sigma model for AFM systems, and the destabilization of the AFM state by quantum fluctuations, with the emergence of an extra "topological Θ-term" in the effective action.

AFM SPIN WAVES: The theory of AFM spin waves is an expansion of small oscillations, ie., of quantum fluctuations, about the classical Néel state. As such we are holding our nose and pretending that the problem just discussed does not exist. Now since the ratio of the quantum fluctuation energy to the classical energy decreases as S increases, we can hope that this tactic may work for large S, (and for large z); we shall see later that this is so.

Let's first quickly do things using a Holstein-Primakoff technique. We generalize (114) to cover 2 sublattices, so that

$$\left.\begin{aligned} S_{jA}^{+} &= (2S)^{1/2}\left[1-\frac{1}{2S}a_j^{+}a_j\right]^{1/2}a_j \\ S_{jB}^{+} &= (2S)^{1/2}\,b_j^{+}\left[1-\frac{1}{2S}b_j^{+}b_j\right]^{1/2} \end{aligned}\right\} \qquad (173)$$

and

$$\left.\begin{aligned} S_{jA}^{z} &= S - a_j^{+}a_j \\ S_{jB}^{z} &= b_j^{+}b_j - S \end{aligned}\right\} \qquad (174)$$

where the swapping between S^+ and S^- in the defns of these operators for the 2 sublattices, and between $\pm S$ for the defn of S^z, comes because the spins are oppositely oriented in the 2 cases.

We will choose as our Hamiltonian a system with both single ion anisotropy and an external field, so that

$$\mathcal{H} = -\gamma \underline{H}_0 \cdot \sum_{j=1}^{N} \underline{S}_j - \sum_{i\neq j}\sum J_{ij}\,\underline{S}_i\cdot\underline{S}_j - \tfrac{1}{2}\sum_j K_2\,(S_j^z)^2 \qquad (175)$$

$$(J_{ij} < 0\,; \quad K_2 > 0)$$

so that we have "easy-axis" anisotropy; we can rewrite this as

$$\mathcal{H} = -\gamma \underline{H}_0 \cdot \Big(\sum_{i\in A}\underline{S}_i + \sum_{j\in B}\underline{S}_j\Big) - \sum_{i\in A}\sum_{j\in B} J_{ij}\,\underline{S}_i\cdot\underline{S}_j - \tfrac{1}{2}K_2\Big(\sum_{i\in A}(S_i^z)^2 + \sum_{j\in B}(S_j^z)^2\Big) \qquad (176)$$

with separate sums over the 2 sublattices A and B; we notice that we are now using the classical Néel state as a point of departure. Now, with a bit of algebra, we can rewrite this Hamiltonian as

$$\left.\begin{aligned}\mathcal{H} = \tilde{E}_0 + \gamma H_0 \Big(\sum_{i\in A} a_i^\dagger a_i - \sum_{j\in B} b_j^\dagger b_j\Big) - \sum_{i\in A}\sum_{j\in B} S J_{ij}\,(a_i^\dagger a_i + b_j^\dagger b_j + a_i b_j + a_i^\dagger b_j^\dagger) \\ + K_2\Big(\sum_{i\in A} a_i^\dagger a_i + \sum_{j\in B} b_j^\dagger b_j\Big)\end{aligned}\right\} \quad (177)$$

and again, it is useful to Fourier transform here, now writing

$$a_q^\dagger = \Big(\frac{2}{N}\Big)^{1/2} \sum_{i\in A} e^{i\underline{q}\cdot\underline{r}_i} a_j^\dagger \qquad b_q^\dagger = \Big(\frac{2}{N}\Big)^{1/2} \sum_j e^{-i\underline{q}\cdot\underline{r}_j} b_j^\dagger \quad (178)$$

and defining the form factor $\gamma_{\underline{q}}$ via

$$z J_0 \gamma_q = \sum_{j\in B} J_{ij}\, e^{-i\underline{q}\cdot(\underline{r}_i - \underline{r}_j)} \quad (179)$$

so that we have

$$\mathcal{H} = \tilde{E}_0 - 2J_0 z S \sum_{\underline{q}} \left[A_0 a_q^\dagger a_q + B_0 b_q^\dagger b_q + \gamma_q (a_{\underline{q}} b_{\underline{q}} + a_{\underline{q}}^\dagger b_{\underline{q}}^\dagger)\right] \quad (180)$$

and we have defined the following constants:

$$\left.\begin{aligned}\tilde{E}_0 &= \tfrac{1}{2}(J_0 z - K_2) N S^2 \\ A_0 &= 1 + \frac{1}{J_0 z S}(K_2 S + \gamma H_0) \\ B_0 &= 1 + \frac{1}{J_0 z S}(K_2 S - \gamma H_0)\end{aligned}\right\} \quad (181)$$

Now (180) is a quadratic form in the space of operators $a_q, a_q^\dagger, b_q, b_q^\dagger$; we need to find a complex rotation in this 4-d space to diagonalize this form. This is the Bogoliubov transformation*

$$\left.\begin{aligned} a_q &= \alpha_q \cosh\theta_q - \beta_q^\dagger \sinh\theta_q & b_q &= \beta_q \cosh\theta_q - \alpha_q^\dagger \sinh\theta_q \\ a_q^\dagger &= \alpha_q^\dagger \cosh\theta_q - \beta_q \sinh\theta_q & b_q^\dagger &= \beta_q^\dagger \cosh\theta_q - \alpha_q \sinh\theta_q \end{aligned}\right\} \quad (182)$$

so that

$$\tanh 2\theta_q = \frac{\gamma_q}{1 + (K_2/J_0 z)} \quad (183)$$

and we see that θ_q is a rotation in a "Lorentzian space" of the original operators

---

* Again, due to Holstein-Primakoff (1940); Bogoliubov was in 1947 and 1958.

The final form for the Hamiltonian is then

$$\mathcal{H} = E_0 + \hbar \sum_q \left[\omega_q^+ \left(\alpha_q^+ \alpha_q + \tfrac{1}{2}\right) + \omega_q^- \left(\beta_q^+ \beta_q + \tfrac{1}{2}\right) \right] \tag{184}$$

where the eigenfrequencies are given by

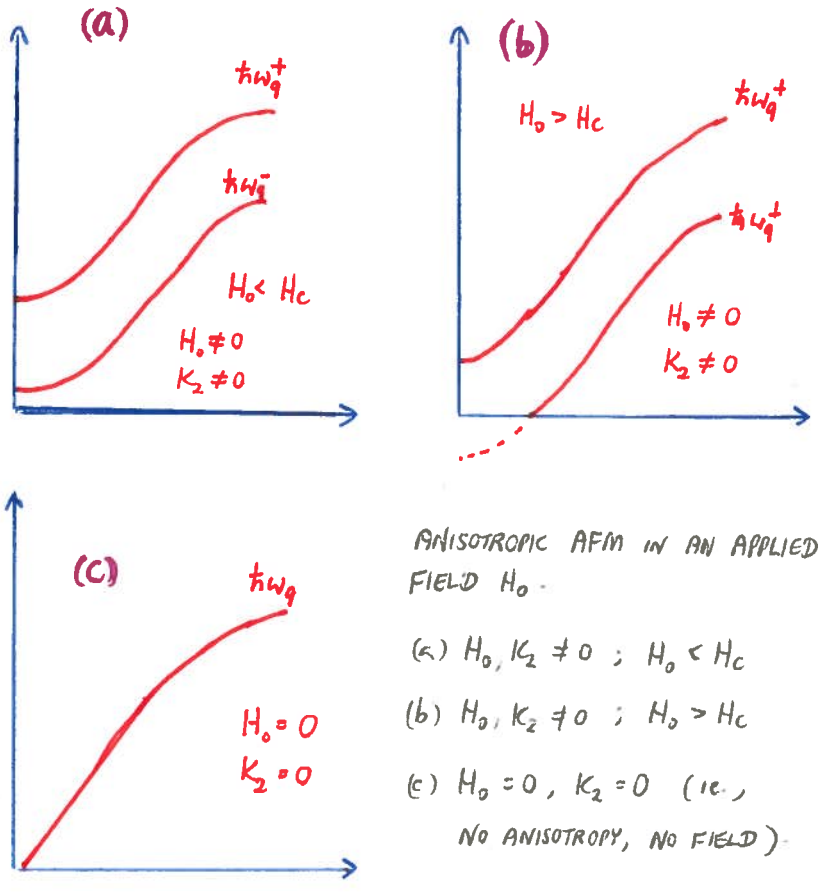
$$\left.\begin{aligned} \hbar\omega_q^\pm &= \left(\Delta_0 + (1-\gamma_q^2) J_0^2 z^2 S^2\right)^{1/2} \pm \gamma H_0 \\ &\xrightarrow[\text{small } q]{} \left(\Delta_0 + c_0^2 q^2\right)^{1/2} \pm \gamma H_0 \end{aligned}\right\} \tag{185}$$

where the "gap" Δ_0 is

$$\Delta_0 = \left(K_2^2 + 2 J_0 z K_2\right)^{1/2} S \tag{186}$$

and the constant c_0 depends on the detailed form of γ_q. Notice what happens here when the anisotropy $K_2 \to 0$; then we have < LINEAR dispersion for the AFM spin waves, instead of the quadratic dispersion found in the FM case, ie., we have

$$\text{AFM spin waves:} \qquad \hbar\omega_q^\pm \xrightarrow[K_2 = 0]{} c_0 |q| \pm \gamma H_0 \tag{187}$$



ANISOTROPIC AFM IN AN APPLIED FIELD H_0.

(a) $H_0, K_2 \neq 0$; $H_0 < H_c$

(b) $H_0, K_2 \neq 0$; $H_0 > H_c$

(c) $H_0 = 0$, $K_2 = 0$ (ie., NO ANISOTROPY, NO FIELD)

Now there are some subtleties associated with these results. The first is to do with the sign of ω_q^-. We notice that if there is no anisotropy, then when the applied field $H_0 = 0$, the spectrum is linear and gapless (see Fig (c) in the figure). However if $H_0 \neq 0$, then the energy $\hbar\omega_q^-$ becomes NEGATIVE when $|q|$ is less than a critical value.

The same is true even when $K_2 \neq 0$. Thus for sufficiently small H_0, the anisotropy stabilizes both ω_q^+ and ω_q^-, and keeps them both gapped and positive. But when H_0 exceeds a critical field H_c, the same thing happens; ie., ω_q^- goes negative for small momentum. This critical field is given by

$$H_c^2 = (H_J + H_k) H_{ic} = (\Delta_0/\gamma)^2 \tag{188}$$

where the 2 fields H_J and H_k are given by

$$\left.\begin{aligned} \gamma H_J &= |J| z S \\ \gamma H_k &= K_2 S \end{aligned}\right\} \quad (189)$$

and are commonly referred to as the exchange & anisotropy fields respectively; we can rewrite (185) as

$$\left.\begin{aligned} \hbar\omega_q^{\pm} &= \gamma\left(H_k(H_J+H_k) + (1-\gamma_q^2)H_J^2\right)^{1/2} \pm \gamma H_0 \\ &= \gamma\left(H_c^2 + (1-\gamma_q^2)H_J^2\right)^{1/2} \pm \gamma H_0 \end{aligned}\right\} \quad (190)$$

Now, although we will not take the time to demonstrate it here, this instability is one in which the spins "cant over" in response to the field. To see what is meant by this, let's now rewrite things in a slightly different form, which brings out the fact that because pairs of spins are involved in the AFM ordering, we actually have 2 independent vectorial degrees of freedom involved in the low-energy spin dynamics, and the best way to describe these is not using the 2 individual sublattice magnetizations, which are strongly coupled, but instead by writing $\underline{S}_j(t)$ in the form (here $\alpha_j = \pm$ identifies sublattice A or B):

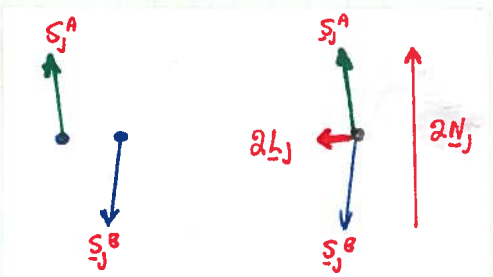
$$\underline{S}_j(t) = S\alpha_j\left[1 - |\underline{m}_j(t)|^2\right]^{1/2} \underline{l}_j(t) + S\underline{m}_j(t) \quad (191)$$

where now the Néel vector $S\underline{l}(t)$ and the canted magnetization vector $S\underline{m}(t)$ represent "difference and sum variables for the 2 sublattices. Note: in the case where the 2 sublattices can be uniquely identified and we can pair off spin $\underline{S}_j^A$ with $\underline{S}_j^B$ in the corresponding mirror lattice, we can also define

$$\left.\begin{aligned} \underline{M}_j &\to \tfrac{1}{2}(\underline{S}_j^A + \underline{S}_j^B) \\ \underline{L}_j &\to \tfrac{1}{2}(\underline{S}_j^A - \underline{S}_j^B) \end{aligned}\right\} \quad (192)$$

However, as defined in (191), $\underline{m}_j(t)$ is zero in the ordered AFM phase where the 2 sublattices are antiparallel; we are not measuring a real magnetization when we define $\underline{M}_j(t)$. Now suppose we start off in zero field, $H_0 = 0$. Then if $K_2 = 0$ also, the Néel vector can point in any direction (provided it is even defined, ie., there is some AFM ordering, which we recall is one of the questions at issue), and we expect $\underline{L}_j = 0$. However as soon as we apply a field, it is energetically advantageous for $\underline{l}_j(t)$ to rotate into the plane perpendicular to $\underline{H}_0$, and a small $\underline{m}_j$ will form, parallel to the field; we note that in general,

$$\left.\begin{aligned} \underline{l}_j \cdot \underline{m}_j &= 0 \\ |\underline{l}_j(t)| &= 1 \end{aligned}\right\} \quad (193)$$



DEFINITION OF THE NÉEL & THE CANTED MAGNETIZATION VECTORS

so that the 2 vectors are always perpendicular. If $K_2 \neq 0$, then in low fields, $\underline{\ell}_j(t)$ will align along the $\pm \hat{z}$ direction, but when $H_0 > H_c$, canting will again occur. In general the phase diagram, & the spin orientations, can become a little complex.

Now we see the physics behind the instability in the spin wave spectrum. In writing our HP operators, we have assumed that the state is ordered, with A and B sublattice spins antiparallel. But the spin wave calculation tells us that for $H_0 > H_c$, there exist disturbances/fluctuations away from this state with negative energy, ie., it is unstable – and now we see why, because the system wants to go to a new canted classical ground state.

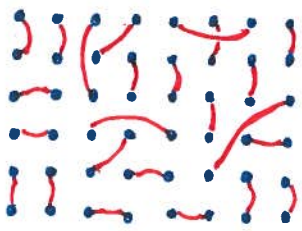
QUANTUM DISORDER & THE HALDANE GAP : All of the above reasoning was basically classical – we started from a classical ground state and then looked at small oscillations about it. But how good is this reasoning?

To answer this, and to illustrate how one may attack the problem using non-perturbative path integral methods, we will reformulate the problem in a different way. We will find a result which is of considerable theoretical interest, because it shows how a topological "θ-term" can appear in the action for the system, and determine the properties.

Before doing this, let's begin with a few more physical questions – this is always a good idea when you are approaching a new topic. We have already seen (cf. eqtns. (169)-(172) that for a pair of spins, there is no reason to expect anything like Néel ordering – so why would we want to use this as a starting point for our theory? This question was posed repeatedly over the years by P.W. Anderson, who argued that one might try a quite different starting point for a ground state, the "RVB" state (where RVB means "Resonating Valence Bond"). The basic idea here is simple – we have seen that for a pair of spins, the singlet state is the lowest energy state for AFM coupling – and of course this state is nothing but the standard valence bond state from elementary chemistry. Let us therefore define the valence bond state between 2 sites i and j as

$$|VB\rangle_{ij} \equiv |\phi_{ij}\rangle = \frac{1}{\sqrt{2}}\left(|\uparrow_i \downarrow_j\rangle - |\downarrow_i \uparrow_j\rangle\right) \qquad (194).$$

Notice that if a site i is in a completely entangled singlet state with site j, then neither i nor j can entangle with another site k (this, in quantum info theory, is given the grand name of the "quantum monogamy theorem"). This means that we can define an RVB configuration like that shown in the figure, in which every site in the lattice is paired off with another one. But of course this configuration is only one of a huge number – without developing any formalism here, we can imagine enumerating all such different configurations, and then summing over them with a different amplitude for each one. One can imagine a huge range of possibilities here, depending on how much one emphasizes short-range or long-range bonds. Thus one extreme has the bonds all between nearest neighbors – this is a "dimerized state", or "valence bond solid". On the



A RESONATING BOND CONFIGURATION.

other hand one can imagine in which much more weight is given to long-range bonds (eg., one in which the weighting is independent of bond length). In this case on expects, as larger & longer bonds come in, to go through a sequence of quite disordered states, and ultimately end up with a Néel-type state, with AFM ordering *

To decide what is the correct ground state (as well as the excited states above it) is very hard in this kind of theory. One can argue that when $\mathcal{Z}$ is large, the large number of AFM couplings involved will stabilize the AFM state (first look at small clusters of spins, then couple these to larger ones, and so on). Certainly it is true that a strong AFM coupling will try and keep neighbouring spins as antiparallel as possible, whether this be in a singlet state, or some other state in which such an AFM correlation is present. But how is one to proceed? And for small $\mathcal{Z}$, can we still make such an argument?

One other way of thinking about this turns out to be very fruitful. We have, when doing fluctuations à la HP, been perturbing around a classical state in which all spins in the system are fixed rigidly in an AFM lattice, in a ground state with long-range order. But we can choose a much more general class of states, in which neighbouring pairs of spins are almost exactly anti-aligned, but where we allow the ordering axis itself to fluctuate over longer length scales. To do this in ordinary field theory, where we must specify the ground state before we start, this is very hard - which state do we choose? But in path integral formulations of the problem, it is easy - we simply assume that the vectors $\underline{\ell}_j(t)$ and $\underline{m}_j(t)$ in (191) are classical, and then the assumption of slow variation of $\underline{\ell}_j(t)$ implies, because of (193), that $\underline{m}_j(t) \ll 1$ everywhere (nearest neighbour spins are almost exactly antiparallel).

Let us therefore rewrite the action in terms of $\underline{\ell}_j(t)$ and $\underline{m}_j(t)$. To keep things simple we assume $H_0 = 0$, $K_2 = 0$, and then we have

$$S_{AFM}[\underline{\ell}_j, \underline{m}_j] = S_B[\{\underline{\ell}_j, \underline{m}_j\}] - \int dt\, \mathcal{H}[\{\underline{\ell}_j, \underline{m}_j\}] \tag{195}$$

where the 1st term (the Berry phase or topological term) is

$$\left.\begin{aligned} S_B &= \hbar S \sum_j \int dt\, \underline{A} \cdot \dot{\underline{n}}_j(t) \\ &= \hbar S \sum_j \int dt \left\{ \underline{A} \cdot (\alpha_j \dot{\underline{\ell}}_j(t) + \dot{\underline{m}}_j(t)) + \dot{\ell}_j^\mu \frac{\partial A_\mu}{\partial \ell^\nu} m^\nu(t) \right\} \end{aligned}\right\} \tag{196}$$

where to get the 2nd expression we use $|\underline{m}(x)| \ll 1$, and expand to 1st order in $\underline{m}(t)$. To deal with the latter term in this expression, we first integrate by parts, and then use $(\nabla \times \underline{A}) \cdot \underline{\ell}_j(t) = 1$ (compare (59)), and then get

$$S_B = \hbar S \sum_j \int dt \left\{ \alpha_j \underline{A} \cdot \dot{\underline{\ell}}_j(t) - (\underline{\ell}_j \times \dot{\underline{\ell}}_j) \cdot \underline{m}_j \right\} \tag{197}$$

in which we now have both $\underline{\ell}_j$ and $\underline{m}_j$; a total time derivative has been dropped

---

* See, e.g., Liang, Doucot, & Anderson, PRL 61, 365 (1988)

here as well.

The Hamiltonian can be handled similarly. We again choose the isotropic nearest-neighbour exchange Hamiltonian, and so we get (cf. (109)):

$$\left.\begin{aligned}\mathcal{H} &= -S^2 J \sum_{\langle ij\rangle} \underline{n}_i \cdot \underline{n}_j \\ &= -JS^2 \sum_{\langle ij\rangle} \left\{ \alpha_i \alpha_j \left[1 - \tfrac{1}{2} (\underline{r}_{ij} \cdot \underline{\nabla} \underline{\ell}_i)^2 \right] + 2 \underline{m}_i^2 \left[1 - \alpha_i \alpha_j \right] \right\} + \cdots \end{aligned}\right\} \quad (198)$$

where we again expand to lowest order in the gradients of $\underline{n}_i$ (and hence lowest order in $\underline{m}_i$).

Now for nearest neighbours, $\alpha_i \alpha_j = -1$ always, so we can now write the Hamiltonian in the continuum approximation (valid because we are looking here at slow variations in space) in the form (dropping constants)

$$\mathcal{H} = -JS^2 \int d^D r \; \tfrac{1}{2} \left[(\nabla \underline{\ell}(x))^2 + 8 \underline{m}^2(x) \right] \quad (199)$$

so that now we can write the generating functional of the system as

$$Z[\underline{j}, \underline{h}] = \int \mathcal{D}\underline{\ell}(x) \int \mathcal{D}\underline{m}(x) \; e^{\frac{i}{\hbar}\left[S_{eff}[\underline{\ell}, \underline{m}] + \int d^D x (\underline{j} \cdot \underline{\ell} + \underline{m} \cdot \underline{h}) \right]} \quad (200)$$

where $\underline{j}(x)$ and $\underline{h}(x)$ are external fields coupling to $\underline{\ell}(x)$ and $\underline{m}(x)$, and the effective action $S_{eff}[\underline{\ell}, \underline{m}]$ is

$$\left.\begin{aligned} S_{eff}[\underline{\ell}, \underline{m}] = \hbar S \int dt \Big\{ & \sum_j \alpha_j \underline{A} \cdot \dot{\underline{\ell}}_j(t) - \int d^D r \; (\underline{\ell} \times \dot{\underline{\ell}}) \cdot \underline{m} \\ & - \tfrac{1}{2} JS^2 \int d^D r \left[(\nabla \underline{\ell})^2 + 8 \underline{m}^2 \right] \Big\} \end{aligned}\right\} \quad (201)$$

where

We now need to integrate over 2 fluctuating variables, $\underline{\ell}(x)$ and $\underline{m}(x)$, which are orthogonal to each other (cf. (193)), so our job seems a little complicated at first, since the 2 fluctuations are coupled via the 2nd term in the Berry phase.

However, we have already seen that the fluctuations $\underline{m}(x)$ are small, so it does not make sense to go beyond lowest order in these. So we can integrate this out by minimizing the effective action S_{eff} under variations of $\underline{m}(x)$; this immediately gives

$$\underline{m}(x) = \frac{1}{8JS} \left(\underline{\ell}(x) \times \dot{\underline{\ell}}(x) \right) \quad (202)$$

We then notice that $(\underline{\ell} \times \dot{\underline{\ell}})^2 = (\partial_t \underline{\ell})^2$, because of the constraint in (193), so finally we have

$$S_{eff}[\underline{\ell}] = \int dt \left\{ \hbar S \sum_j \alpha_j \underline{A} \cdot \dot{\underline{\ell}}_j - \tfrac{1}{2} \rho_s^0 \int d^D x \left[(\nabla \underline{\ell}(x))^2 + \frac{1}{c_0^2} (\partial_t \underline{\ell}(x))^2 \right] \right\} \quad (203)$$

where we have introduced 2 constants, viz., the bare "spin stiffness" ρ_s^0 and the AFM spin wave velocity c_0, given by

$$\left.\begin{aligned} \rho_s^0 &= JS^2 \\ c_0^2 &= 8J^2S^2 \end{aligned}\right\} \quad (204)$$

Now let us see what we have here. The first term in (203) is what is left of the Berry phase - we shall look at it in a moment. The rest of the action is just the action of the non-linear sigma model, which we already introduced previously (cf. eqns. (119) - (127)). As noted already, this model is one of great interest in many different fields; we shall use it as a toy model to study renormalization in the next chapter. The spin stiffness ρ_s^0 is the analogue, for a spin system, of the superfluid density - this just measures the resistance of the spin system to the bending of the spins away from their most favourable state of anti-alignment (for nearest neighbour spins).

The Berry term in (203) has been left for the moment in its lattice form; but we would like to have a continuum form. Now it turns out that we have to do this with care; and the answer depends on the number of spatial dimensions, as follows

1-d AFM Chain: This is the famous case elucidated by Haldane in 1983, leading to the prediction of the "Haldane gap", and of a key difference between integer and ½-integer spin chains*. Suppose we have a function that is varying slowly as we go from one lattice site to the next. Calling this function f_j, we see that

$$\sum_j (-1)^j f_j \qquad \tfrac{1}{2}\int dx\, (\partial f/\partial x) \quad (205)$$

plus terms $\sim O(f''(x), (f'(x))^2, \text{etc})$. Applying this to the Berry phase term in (203), which for a 1-d lattice just has $\alpha_j = (-1)^j$, we get the Haldane result, viz.,

$$\left.\begin{aligned} S_B[\underline{l}] &= \hbar S \int dt \sum_j (-1)^j \underline{A}\cdot\dot{\underline{l}}_j \\ &\to \tfrac{1}{2}\hbar S \int dt \int dx\; \underline{l}\cdot(\partial_x \underline{l} \times \partial_t \underline{l}) \\ &\equiv \tfrac{1}{2}\hbar S \int dt \int dx\; F_{xt}(x,t) \;=\; 2\pi\hbar S\, Q_{xt} \end{aligned}\right\} \quad (206)$$

where the topological term $Q_{\mu\nu}$ is, in D dimensions

$$Q_{\mu\nu} = \frac{1}{4\pi}\int d^D x\, F_{\mu\nu}(x) = \frac{1}{4\pi}\int d^D x\; \underline{l}(x)\cdot(\partial_\mu \underline{l}(x) \times \partial_\nu \underline{l}(x)) \quad (207)$$

(compare eqns. (128) and (129)). Now eqn (206) is an important result, so let's unpack it in several different ways.

- The first way is to use your geometrical intuition. Let's look at the 1st

---

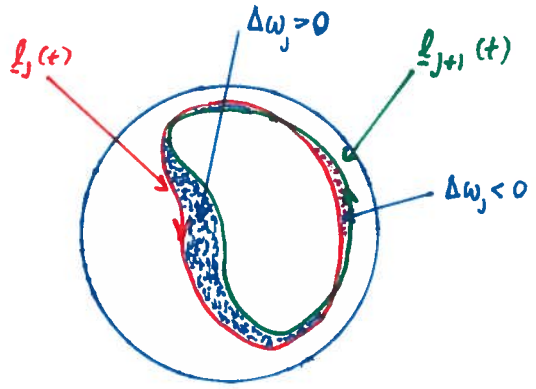
* F.D.M. Haldane, Phys. Rev. Lett. 50, 1153 (1983); and Phys. Lett. 93A, 464 (1983)

line of (206), and ask how to think about it geometrically. Now for each site, we see that the action S_B is measuring the Berry phase / solid angle accumulated by the motion, ie., we have

$$S_B[\underline{\ell}] = \hbar S \sum_j (-1)^j \omega_j[\underline{\ell}_j] \tag{208}$$

and so if we now look at adjacent pairs of $\underline{\ell}$ vectors, we are simply looking at the difference between $\omega_j[\underline{\ell}_j]$ and $\omega_{j+1}[\underline{\ell}_{j+1}]$. To see this geometrically, let's look at 2 closed curves, and define

$$\left.\begin{aligned} \Delta\omega_j &= \omega_j[\underline{\ell}_j] - \omega_{j+1}[\underline{\ell}_{j+1}] \\ &= \int \underline{A} \cdot (d\underline{\ell}_j - d\underline{\ell}_{j+1}) \end{aligned}\right\} \tag{209}$$



CURVES FOR $\underline{\ell}_j(t)$ AND $\underline{\ell}_{j+1}(t)$ ON BLOCH SPHERE. THE DIFFERENCE $\Delta\omega_j = \oint \underline{A}\cdot(d\underline{\ell}_j - d\underline{\ell}_{j+1})$ IS ALSO SHOWN

which for 2 closed curves just gives the picture shown in the figure - we are looking at the difference between the 2 solid angles (between the red and green curves). If we now assume this difference is small, we see that we are integrating the rate of change " of $\omega(x)$ as we go down the 1-d system (in a continuum approximation), ie., we are calculating the integral

$$S_B = \tfrac{1}{2}\hbar S \int dt \int dx \, \frac{d}{dx}\left(\underline{A}(\underline{\ell})\cdot\frac{d}{dt}\underline{\ell}(x,t)\right) \tag{210}$$

where I write $\underline{A}(\underline{\ell})$ to signify that the vector potential $\underline{A}$ is referring to the vector $\underline{\ell}$, and that since $\underline{\ell}\cdot(\nabla\times\underline{A}(\underline{\ell})) = 1$, we can also write that

$$\underline{\ell} \equiv (\nabla_{\ell}\times\underline{A}(\underline{\ell})) \tag{211}$$

(compare eqtn (59)).

- A second way to think about (206) is to subtract off a term which is actually zero, and which allows us to rewrite the action in a more symmetric form. So let us write, from (210), that

$$\left.\begin{aligned} S_B[\underline{\ell}] &= \tfrac{1}{2}\hbar S \int dt \int dx \left\{ \frac{d}{dx}\left(\underline{A}(\underline{\ell})\cdot\frac{d}{dt}\underline{\ell}(x,t)\right) - \frac{d}{dt}\left(\underline{A}(\underline{\ell})\cdot\frac{d}{dx}\underline{\ell}(x,t)\right)\right\} \\ &= \tfrac{1}{2}\hbar S \int dt \int dx \left[\left(\partial_x \underline{A}(\ell)\cdot\partial_t\underline{\ell}\right) - \left(\partial_t\underline{A}_{\ell}\cdot\partial_x\underline{\ell}\right)\right] \\ &= \tfrac{1}{2}\hbar S \int dt \int dx \left[(\nabla_{\ell}\times\underline{A}(\underline{\ell}))\cdot(\partial_x\underline{\ell}\times\partial_t\underline{\ell})\right] \\ &= \tfrac{1}{2}\hbar S \int dt \int dx \left[\underline{\ell}\cdot(\partial_x\underline{\ell}\times\partial_t\underline{\ell})\right] \end{aligned}\right\} \tag{212}$$

which reproduces (206). Then finally

- We can simply subtract off a total time derivative from our initial formula, as follows:

$$\left.\begin{aligned} S_B[\underline{\ell}] &= \tfrac{1}{2}\hbar S \int dt \sum_j (-1)^j \underline{A}\cdot\dot{\underline{\ell}}_j \\ &= \tfrac{1}{2}\hbar S \int dt \int dx \left[\frac{d}{dx}\left(\underline{A}(\ell)\cdot\frac{d}{dt}\underline{\ell}(x,t)\right) - \frac{d}{dt}\left(\underline{A}(\underline{\ell})\cdot\frac{d}{dx}\underline{\ell}(x,t)\right)\right] \end{aligned}\right\} \quad (213)$$

where we assume boundary conditions such that $\underline{A}\cdot(d\underline{\ell}/dx)$ is unchanged between the initial and final states. It is then straightforward, using the same manoeuvres as above, to get the final answer in (206).

We can now summarize all of this by saying that for a 1-dimensional AFM coupled spin chain, the action takes the form

$$\left.\begin{aligned} S_{eff}[\underline{\ell}] &= S^{NLS}[\ell] + S^{Top}[\ell] \\ S^{NLS}[\underline{\ell}] &= -\tfrac{1}{2}\rho_s^0 \int dt \int dx \left[(\partial_x \underline{\ell}(x,t))^2 + \tfrac{1}{c_0^2}(\partial_t \underline{\ell}(x,t))^2 \right] \\ S^{Top}[\underline{\ell}] &= \frac{\theta}{4\pi}\int dx \int dt\ \underline{\ell}\cdot(\partial_x\underline{\ell}\times\partial_t\underline{\ell}) \\ &\equiv \tfrac{1}{2}\hbar S \int dx \int dt\ \underline{\ell}\cdot(\partial_x\ell\times\partial_t\ell) \equiv \theta Q_{xt} \end{aligned}\right\} \quad (1d) \qquad (214)$$

where the "topological θ-parameter" $\theta = 2\pi\hbar S$.

Now to fully elucidate the properties of this system, we need to know how to treat the non-linear sigma term $S_{NLS}[\underline{\ell}]$; and this we can't do until we have looked more fully at renormalization and the renormalization group. However we can look at the "θ-term", or Berry phase term, in more detail - it is simple to analyse.

Consider first the Pontryagin number Q_{xt}; it is quantized, and in fact we see that the total topological term is just going to be

$$S_{Top}[\underline{\ell}] = n\theta \equiv 2\pi\hbar n S \qquad (215)$$

so that we get a phase in the action given by $2\pi n S$. As we have already seen, this distinguishes critically between integer and ½-integer spins (cf. eqns. (107), (108)); in the case where S is an integer, it makes no difference at all to the physics, since $\exp\{4\pi i S\} = 1$. However for ½-integer spin, we can get contributions in the action of ± 1, coming from even/odd n in (215). When we sum over these, it is not obvious what we will get.

To answer this question, as noted above, we need to deal with the other term $S_{NLS}[\underline{\ell}]$ in the total action. However we can simply state here what is found: For ½-integer spins, the effect of the Berry phase θ-term is to make the system gapless (as one finds using simple spin wave theory). However when S = integer, the system develops a large gap in the spin wave spectrum, even when there is no anisotropy or external field. This gap is called the "Haldane gap";

and we will return to it in the next chapter.

<u>AFM lattice, dimension $d \geq 2$</u> : A very obvious question we can ask is – does the Θ-term exist for higher dimensions? This question was of great interest in the early days of high-T_c superconductivity, until it was discovered that the answer is, at least for most cases, no.

To see this, let's repeat the same sort of analysis as we did for the 1-d case, but now for the (2+1)-dimensional case, ie., for spatial dimension $d = 2$. Then instead of (204) we have

$$S_B[\underline{\ell}] = \hbar S \int dt \sum_{ij} (-1)^{i+j} \underline{A} \cdot \dot{\underline{\ell}}_{ij}(t) \tag{216}$$

where now i and j refer to lattice sites measured along the x and y directions. In the same way as we went to the continuum limit in (205) for a chain, we can do the same in 2-d for a lattice function f_{ij}:

$$\sum_i \sum_j (-1)^{i+j} f_{ij} \longrightarrow \tfrac{1}{4} \int dx \int dy \left(\frac{\partial^2 f}{\partial x \, \partial y} \right) \tag{217}$$

and so then (216) transforms in this limit to

$$\left.\begin{aligned} S_B[\underline{\ell}] &\longrightarrow \tfrac{1}{2}\hbar S \int dt \int dx \int dy \; \partial_y \left[\partial_x \left(\underline{A} \cdot \frac{d}{dt} \underline{\ell}(x,y;t) \right) \right] \\ &= 2\pi\hbar S \int dy \; \partial_y \left(Q_{xt}(y) \right) \end{aligned}\right\} \tag{218}$$

where the quantity $Q_{xt}(y)$ is the topological term as a function of x and t, fixed at a specific y:

$$Q_{xt}(y) = \frac{1}{4\pi} \int dx \int dx \; \underline{\ell}(x,y;t) \cdot \left(\partial_x \underline{\ell}(x,y;t) \times \partial_t \underline{\ell}(x,y;t) \right) \tag{219}$$

Now, however, the problem appears. For a given fixed y, $Q_{xt}(y)$ is a fixed topological charge; and changing y will not change this. Thus – this is the standard argument – $\partial_y Q_{xt}(y) = 0$. It then follows that for a 2-d lattice (and also for $d = 2$), the Berry term is zero. We will return to this point later, because this argument can fail if there are singularities in the $\underline{\ell}(r,t)$ field.

It turns out that all of this theory agrees well with experiment, once one has handled the non-linear sigma term (see next chapter). We therefore conclude that even in 1 dimension, we can describe the AFM-coupled spin system as though it has short-range order; but there may well be no long-range order (a conclusion which will be demonstrated in the next chapter). The effect of the topological term is profound, because it lifts the gap in what would have otherwise been a Goldstone boson mode. This conclusion doesn't yet follow from what we've done – again, it requires renormalization group methods.

B.6.2 : FAST & SLOW VARIABLES

A central element in the formulation of all physical theories is the separation of physical processes into different scales - both time scales and length scales. In what follows, we emphasize that ALL theories in physics involve this - that there is no such thing as a genuinely "microscopic" theory, and that all Hamiltonians or Lagrangians are "effective".

The standard mantra in physics, which is formulated in the ideas of "renormalization", and "integrating out" degrees of freedom, is that when we write down an effective Lagrangian or effective Hamiltonian, we have already performed a kind of "coarse graining" of the system, in which very fast and very small-scale processes and structures have been eliminated from the description, and incorporated into an effective theory into which these fast, small objects have been absorbed into larger and slower entities which now form the basic elements of the theory. To illustrate this, let's consider 2 examples:

(1) A Hydrogen Atom in a Box: Suppose we have a container, an absolutely impermeable box of volume V_0 (assume $V_0 = L_0^3$), containing a single H atom, along with photons in a thermal distribution at temperature T. How then should we describe this system? It is clear that if the temperature satisfies $kT \ll E_0$, where E_0 is the ionization energy, we can write a Hamiltonian

$$H = P^2/2M + \sum_q \hbar\omega_q \left(b_q^\dagger b_q + 1/2\right) + V_{int}\left(\Phi, b_q, b_q^\dagger\right) \quad (220)$$

where $M = M_0 + m_e$ is the combined mass of the proton and electron, P is the momentum of the H atom, b_q and $b_q^\dagger$ are photon operators, and V_{int} is the weak interaction between the photons and the position Φ of the atom.

But what if the photons are at a much higher temperature, say 10^5 K? Then we must write

$$\mathcal{H} = \frac{P_N^2}{2M_0} + \frac{P_e^2}{2m_e} + \frac{e^2}{|\underline{R}_N - \underline{r}_e|} + \sum_q \hbar\omega_q \left(b_q^\dagger b_q + 1/2\right) + \tilde{V}_{int}\left(\underline{R}_N, \underline{r}_e, b_q, b_q^\dagger\right) \quad (221)$$

where we are still working with a non-relativistic Hamiltonian, but now with Coulomb interactions between electron and proton, and the much stronger couplings between these 2 charged particles and the photons. Notice that even the vacua of these 2 Hamiltonians are different; in the first case we have a photon vacuum along with the vacuum for a single uncharged particle, whereas in the second, the vacuum is that for 2 charged particles plus photons.

If we continue to raise the energy we can go up to $kT \sim m_e c^2$, i.e., $kT \sim 5 \times 10^9$ K. At this point we must switch to a Lagrangian

$$L(x) = \bar{\psi}_N(x)\left[i\gamma^\mu D_\mu - M_0\right]\psi_N(x) + \bar{\psi}_e\left[i\gamma^\mu D_\mu - m_e\right]\psi_e + F_{\mu\nu}(x) F^{\mu\nu}(x) \quad (222)$$

and, to fix the constraint that the net electronic charge in the system is unity (and likewise for the proton), we can add Lagrange multipliers to fix these, in the action. Now the vacuum is the full QED vacuum, with both nucleon and electron fields, and it will show strong vacuum fluctuations, with electron-positron pairs appearing all over the place.

One can of course continue this to even higher T, but the point is clear; each of these descriptions involves the assumption of some set of building blocks with effective couplings between them, and all more microscopic processes are incorporated implicitly into these.

(ii) A simple Fluid: Consider a gas of, eg., ^{3}He atoms. At a temperature of, eg., 100 K, we can describe this by a simple effective Hamiltonian of form

$$\mathcal{H} = \sum_j \frac{\underline{P}_j^2}{2M_3} + \sum_{i \neq j} V(\underline{r}_i - \underline{r}_j) \qquad (\text{fast}) \tag{223}$$

where M_3 is the mass of the ^{3}He atoms, and $V(\underline{r}_i - \underline{r}_j)$ the non-relativistic effective interaction between them (with a short-range strong repulsion, and long-range weak attraction).

However if we are only interested in the "long-wavelength" properties of the system at this temperature — ie., those occurring at length scales $\gg n^{-1/3}$, where n is the particle density, and timescales $\gg \tau_0$ (the mean free time between collisions) then we can go over to the continuum description embodied in the Navier-Stokes equation of motion, viz.,

$$\frac{d}{dt}\underline{v}(\underline{r},t) + \nu \underline{\nabla}^2 \underline{v}(\underline{r},t) = -\underline{\nabla} p(\underline{r},t) \qquad (\text{slow}) \tag{224}$$

where $p(\underline{r},t)$ is the pressure, and ν the dynamical velocity; this is known as a "hydrodynamic" description, and we will be discussing it in much more detail later on.

If we now cool the ^{3}He to, say, 10 mK, it will liquefy (and solidify if the pressure is high enough) and then we go over the Landau Fermi Liquid Theory description, with an effective Hamiltonian

$$\mathcal{H} = \sum_{\underline{k}\sigma} \varepsilon_{\underline{k}\sigma} c^+_{\underline{k}\sigma} c_{\underline{k}\sigma} + \sum_{\underline{k}\underline{k}'\underline{q}\sigma\sigma'} f^{\sigma\sigma'}_{\underline{k}\underline{k}'} c^+_{\underline{k}+\underline{q},\sigma} c_{\underline{k}'\sigma'} c^+_{\underline{k}'-\underline{q},\sigma} c_{\underline{k}\sigma} \tag{225}$$

involving quasiparticle operators $c^+_{\underline{k}\sigma}$, $c_{\underline{k}\sigma}$, a quasiparticle energy $\varepsilon_{\underline{k}\sigma}$, and effective FLT interactions $f^{\sigma\sigma'}_{\underline{k}\underline{k}'}$. If we go to still lower temperatures, we go through a superfluid transition, and the effective Hamiltonian becomes very complicated indeed. Note again that the vacua of these different descriptions are all quite different from each other.

In what follows we will discuss various techniques for separating out the fast and slow processes in a DFT, and extracting the key information about each. We note that we will be focussing on a single system at a given temperature, and switching between two energy/length scales — this is different from

what happens when we change the temperature, which will change completely the nature of the system itself.

B.6.2 (a) GRADIENT EXPANSIONS

Let us begin by recalling that a semi-classical expansion about the classical theory, introduced in section B.6.1 (b) and in typical WKB form in eqtns (67)-(73), is an expansion in powers of $\hbar$; but it also involves an expansion in spatial and temporal gradients, as we see from the first few terms in the expansion, and from the general recursion relation in (69). One sees this also in the lowest "eikonal" term in the expansion of the Schrodinger eqtn; if we write

$$\psi(r,t) = A(r,t)\, e^{iS/\hbar} \tag{226}$$

and substitute into the Schrodinger eqtn., viz.,

$$\left[-\frac{\hbar^2}{2m} + V(r)\right]\psi = i\hbar\,\partial_t\,\psi \tag{227}$$

we find, assuming that $A(r,t)$ is slowly-varying, the lowest-order eqtn (of order $\hbar^0$, ie., independent of $\hbar$):

$$\frac{1}{2m}\left(\nabla S(r,t)\right)^2 + V(r) = -\partial_t S(r,t) \tag{228}$$

which is just the Hamilton-Jacobi eqtn for $S(r,t)$, the action. One can continue this along the lines given in eqtns (67)-(69).

Let's now start off by doing some explicit gradient expansions, in several different ways.

(i) GRADIENT EXPANSION FOR DRIVEN SCALAR FIELD

We have already discussed what happens when a scalar field (or for that matter, an ordinary particle in, eg., a harmonic well) is subject to an external field - whether this field just be some applied force, or whether it be some time- and space-varying potential. Thus, eg., we could pick an action

$$S = \frac{1}{2}\int d^3x\left[\left(\partial_\mu\phi\,\partial^\mu\phi - m^2\phi^2\right) + V(x)\,\phi^2(x)\right] \tag{229}$$

in which the external field acts on $\phi(x)$ in addition to the constant mass. If $V(x)$ is just treated as some external field, then we leave everything at the end as a functional of $V(x)$, so that, eg., the 2-point propagator takes the form $\mathcal{G}(x,x'|V)$, and the partition function / generating functional can be

written as a functional $Z[J,V]$ of both the potential $V(x)$ and the applied field $J(x)$.

The idea now is to pick out the long-wavelength, slow parts of this, in powers not of the coupling constant but of the spacetime gradients. This means that the results ought to be valid away from weak coupling, if we do things properly.

To illustrate this let's start with a problem we already understand really well, which is a driven free scalar field, with the driving coming from a current source, giving

$$S[J] = \int d^4x \, \tfrac{1}{2}\left(\partial_\mu\phi\partial^\mu\phi - m^2\phi^2\right) - J(x)\phi(x) \tag{230}$$

so that we deal with

$$Z[J] = \int \mathcal{D}\phi(x)\, e^{i S[J]} \tag{231}$$

We know, of course, how to deal with this problem exactly; but let's look at it now in a gradient expansion. The 1-particle propagator satisfies the eqtn. of motion

$$\left(\partial^2 - m^2 + J(x)\right) G(x,x'|J) = -\delta(x-x') \tag{232}$$

Now if $J(x)$ is considered to be slowly varying, we expect $G(x,x'|J)$ to be a slowly-varying function of the "centre of mass" coordinate $\tfrac{1}{2}(x+x')$, but not necessarily of $x-x'$ (to convince yourself of this, consider the simple example of an oscillator, subject to a force $J(t)$; if $J(t)$ is a constant, all we do is shift the oscillator, and if it is slowly-varying compared to the natural frequency ω_0 of the oscillator, then it will hardly disturb the oscillator; but the oscillator will still have natural oscillations on a timescale $2\pi/\omega_0$, which will be reflected in the Green function $G(t,t'|J(t))$).

Thus we see that if we have the condition

$$\left|\frac{\partial_\mu J}{J}\right| \ll m. \tag{233}$$

we can assume slow variation. Let's now Fourier transform with respect to $x-x'$, so that

$$G_k(x|J) = \int d^4(x-x')\, G(x,x'|J)\, e^{-ik\cdot(x-x')} \tag{234}$$

& so that the eqtn of motion now reads

$$\left[(k_\mu + i\partial_\mu)(k^\mu + i\partial^\mu) + m^2 - J(x)\right] G_k(x|J) = 1 \tag{235}$$

where for simplicity we write things in terms of x rather than $\tfrac{1}{2}(x+x')$. We are now going to deal with the formal solution of this eqtn in 2 different ways. We have formally that

$$G_k(x) = \frac{1}{\left(k^2 - \partial_x^2 + 2ik_\mu\partial_x^\mu + m^2 - J(x)\right)} \tag{236}$$

and the idea is to now use the slow variation of $J(x)$, and of the gradient operators in this result, as the basis for an expansion. How we do this depends on what we choose for our unperturbed result, and what we choose as the expansion parameter.

(a) **FLUCTUATION EXPANSION** : Suppose we treat the gradient operators as the perturbation, and so write the propagator eqtn of motion as

$$\left[G_0^{-1}(k,x|J) - W_k(\partial_x^\mu) \right] \mathcal{G}_k(x|J) = 1 \tag{237}$$

where we have

$$G_0(k,x|J) = \frac{1}{k^2+m^2-J(x)} \tag{238}$$

and

$$W_k(\partial_\mu) = \partial_x^2 - 2ik_\mu \partial_x^\mu \tag{239}$$

where we have written ∂_x^μ and ∂_x^2 to explicitly indicate that the operator ∂_μ is acting on the variable x (ie, the variable upon which things depend slowly).

We can now write a Dyson expansion for the eqtn of motion; we have

$$\left.\begin{aligned} \mathcal{G}_k(x|J) &= \left[G_0^{-1}(k,x|J) - W_k(\partial_\mu) \right]^{-1} \\ &= G_0(k,x|J) + G_0(k,x|J)\, W_k(\partial_\mu)\, G_0(k,x|J) + \cdots \end{aligned}\right\} \tag{240}$$

Note that this form of the expansion is only useful if we have a simple form for $G_0(k,x|J)$. To see how this works let's determine the lowest corrections to G_0 in this expansion. We have

$$\left.\begin{aligned} \delta^{(1)}\mathcal{G}_k(x|J) &= G_0 W G_0 \\ &= \frac{1}{k^2+m^2-J(x)} \left(\partial^2 - 2ik_\mu\partial^\mu\right) \frac{1}{k^2+m^2-J(x)} \\ &= G_0^3(k,x|J)\left[\left(\partial^2 - 2ik_\mu\partial^\mu\right)J(x)\right] \quad \sim \; -2i\,G_0^3(k,x|J)\,k_\mu\partial^\mu J(x) \end{aligned}\right\} \tag{241}$$

and then for the 2nd correction

$$\left.\begin{aligned} \delta^{(2)}\mathcal{G}_k(x|J) &= G_0 W G_0 W G_0 \\ &= G_0^3(k,x|J)\left[\left(\partial^2 - 2ik_\mu\partial^\mu\right)^2 J(x)\right] + 3\,G_0^5(k,x|J)\left[\left(\partial^2 - 2ik_\mu\partial^\mu\right)J(x)\right] \\ &\sim \; -4\,G_0^3(k,x|J)\,k_\mu\partial^\mu k_\nu\partial^\nu J(x) - 6i\,G_0^5(k,x|J)\,k_\mu\partial^\mu J(x) \end{aligned}\right\} \tag{242}$$

where in the approximate result we assume that $|k_\mu\partial^\mu J(x)| \gg \partial^2 J(x)$, ie, that the important values of k extend over a much wider range than the Fourier transform of $J(x)$ (because things vary more rapidly with $x-x'$ than with x).

If we wish we can further Fourier transform $\mathcal{G}_k(x|J)$, so that everything is now

written in terms of the Fourier transform of $J(x)$, ie., in terms of

$$J_q = \int d^4x\, e^{-iqx} J(x) \tag{243}$$

$$\mathcal{G}_{kq}[J_q] = \int d^4x\, e^{-iqx}\, \mathcal{G}_k(x|J) \tag{244}$$

We then have

$$\left[(k^2+q^2+2k_\mu q^\mu) + m^2 - J_q\right] \mathcal{G}_{kq}[J_q] = 1 \tag{245}$$

and we can then calculate everything from this.

We can of course represent this Dyson eqtn diagrammatically, as shown in the figure; the extend perturbing fields are now the gradient operators

so that we can think of the extend gradient operator $W_{k_\mu}(\partial^\mu)$ as a self-energy operator.

(b) FRADKIN-STYLE WKB EXPANSION

: A somewhat more obvious tactic to employ in making a gradient expansion of a function like $\mathcal{G}_k(x|J)$ is to separate off everything but the free field term, ie., write

$$\left[G_0^{-1}(k) - U_k(\partial_\mu, J(x))\right] \mathcal{G}_k(x|J) = 1 \tag{246}$$

where

$$G_0(k) = \frac{1}{k^2+m^2} \tag{247}$$

and

$$U_k(\partial_\mu, J(x)) = \partial_x^2 - 2ik_\mu\partial^\mu + J(x) \tag{248}$$

is the perturbation. Now there is a nice way to represent the gradient expansion which makes it look much more like the WKB expansion we discussed in eqtns. (67)-(73); this uses what is usually called the "Schwinger proper time" or "5th-time" representation (although it is actually due to V. Fock), in a form discussed by Fradkin. We begin by noting we can write the Feynman propagator as

$$G_0(k) = \frac{1}{k^2+m^2-i\delta} = i\int_0^\infty ds\, e^{-is(k^2+m^2-i\delta)} \equiv i\int_0^\infty ds\, G_0(k,s) \tag{249}$$

and now write the total Green function in the form

$$\mathcal{G}_{ic}(x|J) = i\int_0^\infty ds\, e^{i(k^2+m^2-i\delta)} \mathcal{F}_k(x,s|J) \tag{250}$$

where the slowly-varying effects are all encapsulated in the functional $\mathcal{F}_k(x,s|J)$. The parameter s is called the proper time, and we see why this is so if we note that $G_0(k,s)$ satisfies the differential eqtn

$$(k^2+m^2-i\delta)\, G_0(k,s) = i\partial_s G_0(k,s) \tag{251}$$

We can now think of the functional $\mathcal{F}_k(x,s|J)$ as a weighting functional over the free particle propagator; note that when $J(x) = 0$, $\mathcal{F}_k(x,s|J) = 1$.

It is now time to put the factors of $\hbar$ back into all of these eqtns; if we do so, we find that we must replace $U_k(\partial_\mu, J(x))$ by

$$U_k(\partial_\mu, J(x)) = \partial_x^2 - 2ik_\mu\partial^\mu + \hbar J(x) \tag{252}$$

Let us now, in the spirit of WKB, write $\mathcal{F}_k(x,s|J)$ in the form

$$\mathcal{F}_k(x,s|J) = e^{\Phi(x)/\hbar} = e^{1/\hbar \sum_n \hbar^n \psi_n(x)} \tag{253}$$

and then substitute this into the eqtn of motion (compare the substitution of (68) into (67). We find the following differential eqtn for $\psi_n(k,x;s|J)$:

$$(\partial^2 + 2ik_\mu\partial^\mu)\psi_n + \sum_{r=1}^{n-1} \partial_\mu\psi_r\,\partial^\mu\psi_{n-r} + J(x)\,\delta_{n1} = -i\partial_s\psi_n \tag{254}$$

which should be compared with eqtn. (69); note that $\psi_0 = 0$.

We then find the leading correction $\psi_1(k,x;s|J)$ from the differential eqtn.

$$(\partial^2 + 2ik_\mu\partial^\mu + i\partial_s)\,\psi_1(k,x;s|J) = -J(x) \tag{255}$$

This is the same kind of eqtn we run into when trying to solve for any function defined in terms of a Green function with a source $J(x)$; you should easily find that

$$\psi_1(k,x;s|J) = i\int_0^s ds' \sum_q e^{iqx}\,\phi_1(q,x;s'|J) \tag{256}$$

with

$$\phi_1(qx;s'|J) = J_q\, e^{-is'(q^2-2k_\mu q^\mu)} \tag{257}$$

and J_q the Fourier transform of $J(x)$ defined above. The higher terms are given by more complicated forms. The general form for the higher terms can be found, if one wishes, by either using the set of coupled eqtns

in (254), or by rewriting it, using the result we found for ψ_1 in (256), in the form

$$\psi_n(k,x;s|J) = i\int_0^s ds'\, e^{i(s-s')(\partial^2 - 2ik_\mu\partial^\mu)} \sum_{r=1}^{n-1} \partial_\mu \psi_r(k,x;s|J)\,\partial^\mu \psi_{n-r}(k,x;s'|J) \quad (258)$$

Now these results are a little academic, in that we already know how to write a solution to this problem – it is just the well-worn Gaussian result that we've been using since section B.1. Recall that we have

$$Z_0[J] = e^{-\frac{i}{2\hbar}\int d^4x \int d^4x'\, J(x)\Delta_F(x-x')J(x')} \quad (259)$$

and the propagator $G_2(x,x'|J)$ is (cf. eqtn B.1 (67)):

$$G_2(x,x'|J) = -\hbar^2 \frac{\delta^2 Z_0[J]}{\delta J(x)\,\delta J(x')} \quad (260)$$

where now we do <u>not</u> set $J(x) = 0$ after taking the 2nd functional derivative; we have

$$\left.\begin{aligned} G_2(x,x'|J) = {} & \left[\int d^4y \int d^4y'\, \Delta_F(x-y)J(y)\Delta_F(x'-y')J(y') - \Delta_F(x-x')\right] \\ & \times e^{-\frac{i}{2\hbar}\int d^4x \int d^4x'\, J(x)\Delta_F(x-x')J(x')} \end{aligned}\right\} \quad (261)$$

It is an interesting exercise to expand this result directly in a gradient expansion, and compare with what we have found above.

The real use of gradient expansions comes in when we have more than one field. This can happen either when we are dealing with a coupled field system (which we will come to presently, in the form of QED), or when we split a single field into 2 components, viz., a slowly-varying and a rapidly-varying one. This tactic forms the basis of the renormalization and renormalization group approaches (next chapter); it is also at the basis of the "background field method", to which we come very soon.

(ii) GRADIENT EXPANSION for Q.E.D.

We now turn to a real physical theory, where the use of long-wavelength analyses has been of key importance over a long period of time. The best-known problem in QED of this kind is that of infra-red divergences and the contribution of soft photons to electron-photon scattering – this is usually known as the "Bremsstrahlung" problem, where it refers to the deceleration of electrons or other charged particles by photon emission, when they are in the field of some other charged object which changes their trajectory (a charged particle moving in free space cannot emit photons – this is kinematically impossible). It is process

of key importance in nuclear physics, astrophysics, plasma physics, and solid-state physics. The photons are emitted in all directions, with intensity $\propto m^{-4}$ (for radiation perpendicular to the motion of the particle, of mass m), and $\propto m^{-6}$ (for radiation parallel to the motion. This means that light charged particles like electrons lose energy very quickly to this process.

The theoretical problems arise when one tries to analyze this process perturbatively in QED, and it quickly becomes apparent that the process is non-perturbative — indeed, as the energy of the emitted photons goes to zero, the number of photons carrying emitted energy diverges, and one must sum diagrams to infinite order to properly understand it. This phenomenon was called the "IR catastrophe" by Bloch & Nordsieck in 1937, who first analyzed it. In solid-state physics in the late 1960's and 1970's it was called the "X-ray edge" catastrophe, or the "Anderson orthogonality" catastrophe, and it figured large in the discussion of the Kondo problem; it was not finally tamed until the development of renormalization group methods.

That the problem is serious is obvious if we consider that it affects not only all scattering processes involving charged particles, but also the calculation of the one-particle propagator $\mathcal{G}_2(x,x')$ for the electrons — one needs to understand the Bremsstrahlung contribution to find the low-energy form for this propagator.

In what follows we will see how to do a gradient expansion on this problem — it turns out that this does the job, and is exact in the low-energy limit.

ONE-PARTICLE PROPAGATOR $\mathcal{G}_2(x,x'|A)$:

If we try to calculate the Bremsstrahlung process and its effect on $\mathcal{G}_2(x,x')$ using a perturbative expansion in the coupling, we find that we cannot get any sensible results without summing diagrams to infinite order — but which diagrams should we sum?

So instead we adopt the following tactic, which we already developed back in section B.3. We calculate $\mathcal{G}(x,x'|A_\mu(z))$, where we have frozen $A_\mu(z)$ to begin with; and then we functionally average $\mathcal{G}_2(x,x'|A)$ over all different configurations of $A_\mu(x)$, with the appropriate weighting factor.

If we followed the development of section B.3, we would first calculate $Z[\eta,\bar\eta;J_\mu]$, and then derive $\mathcal{G}_2(x,x'|A_\mu)$ from this by functional differentiation. But from what we've done above, we see there is no need to do this — we can simply calculate $\mathcal{G}_2(x,x'|A_\mu)$ directly. We have the eqtn. of motion for the fermion propagator, given by

$$\left[\gamma^\mu(i\partial_\mu + eA_\mu) - m\right]\mathcal{G}_2(x,x'|A_\mu) = -\delta(x-x') \tag{262}$$

in the presence of some fixed or "frozen" background field $A_\mu(x)$, which we will later unfreeze. In the early days of QED, when the first attempts were being made to understand the IR divergence problem, this eqtn. was approximated by the simpler eqtn.

$$\left[c^\mu(i\partial_\mu + eA_\mu) - m\right]\mathcal{G}_2(x,x'|A_\mu) = -\delta(x-x') \tag{263}$$

where the c^μ are numbers, instead of the gamma matrices γ^μ in (262), and which

satisfy the relation $$g_{\mu\nu} c^\mu c^\nu = 1 \tag{264}$$

In this case the free particle propagator $G_0(xx')$ obeys the eqtn.

$$(ic^\mu \partial_\mu - m) G_0(xx') = \delta(x-x') \tag{265}$$

with solution $$G_0(x,x') = e^{ik(x-x')} \frac{i\hbar}{c^\mu k_\mu - m - i\delta} \tag{266}$$

instead of

$$(i\gamma^\mu \partial_\mu - m) S_F(x,x') = -\delta(x-x') \tag{267}$$

with solution $\mathcal{G}_0 = i\hbar S_F$ given by

$$\mathcal{G}_0(x-x') = \sum_k e^{-ik(x-x')} \frac{i\hbar}{\gamma^\mu k_\mu - m + i\delta} \tag{268}$$

(cf chapter B.2, eqtns (52)-(55)).

Let's now do the same thing here as we did for the scalar field theory, finding in proper time representation the result for $\mathcal{G}_2(x,x'|A)$, and for its partial Fourier transform $\mathcal{G}_k(x|A)$. Without going through all the details, we see that we can write

$$\left.\begin{aligned} \mathcal{G}_2(x,x'|A) &= i\int_0^\infty ds\ \mathcal{G}_s(xx'|A) \\ &= i\int_0^\infty ds\ e^{is/\hbar[\gamma^\mu(i\partial_\mu + eA_\mu) - m]}\ \delta(x-x') \end{aligned}\right\} \tag{269}$$

so that

$$[\gamma^\mu(i\partial_\mu + eA_\mu) - m]\,\mathcal{G}_s(xx'|A) = -i\hbar\,\partial_s\,\mathcal{G}_s(x,x'|A) \tag{270}$$

with the boundary condition $\mathcal{G}_s(x,x'|A) \longrightarrow \delta(x-x')$, as $s \to 0$.

Now we Fourier transform the relative coordinate as before, ie., write the δ-function in (269) as

$$\delta(x-x') = \sum_k e^{ik(x-x')} \equiv \int \frac{d^4p}{(2\pi\hbar)^4} e^{\frac{ip}{\hbar}(x-x')} \tag{271}$$

so that we can write $\mathcal{G}_s(xx'|A)$ as

$$\mathcal{G}_s(x,x'|A) = \sum_k e^{-ik(x-x')}\,\mathcal{G}_k(x;s|A) \tag{272}$$

and then write $\mathcal{G}_k(x;s|A)$ in the form (cf. eqtns. (250) and (253)):

$$\mathcal{G}_k(x;s|A) = e^{is/\hbar(\gamma^\mu p_\mu - m + i\delta)}\, e^{i\Phi_s(x|A)} \tag{273}$$

where now we put back the usual $i\delta$ factor in the free propagator, and define $p_\mu = \hbar k_\mu$. The phase factor $\Phi_s(x|A)$ in (273) then satisfies a differential eqtn., which for low energy, such that $p^2 \ll m^2$, taken the simplified

form:

$$\gamma^\mu \left(\partial_\mu \Psi_s + e A_\mu \right) = \partial_s \Psi_s \tag{274}$$

with the boundary condition: $\Psi(s=0) = 0$.

The solution to this differential eqtn. is simple; in analogy with the previous discussion (cf. eqtn (255) - (257)) we Fourier transform to q-space, and get

$$\Psi_s(x|A) = \int_0^s ds' \sum_q e^{-iqx} e^{is'p^\mu q_\mu} e p_\mu A^\mu(q) \tag{275}$$

where $A_\mu(q)$ is the Fourier transform of $A_\mu(x)$. Thus we can write out the complete solution for $G_s(x,x'|A)$ in the form (using (275), (273), and (272)):

$$G_2(x,x'|A) = i \sum_k e^{-ik(x-x')} \int_0^\infty ds\, e^{is/\hbar (\gamma^\mu p_\mu - m + i\delta)} e^{ie \int_0^s ds' \sum_q e^{is'q^\mu p_\mu} e^{-iqx} p^\mu A_\mu(q)} \tag{276}$$

Notice 2 things about this solution, in comparison with the discussion of the scalar field. First, we do not need to go to an expansion in powers of $\hbar$ in the phase Ψ_s, and so there is no set of recursive eqtns as in (254). This is because the low-E Dirac eqtn is only linear in derivatives ∂_μ, instead of quadratic like scalar field theory (or the Schrödinger eqtn). Thus (276) is a closed solution. The 2nd thing we notice is that the phase factor $\Psi_s(x|A)$ does not depend on the relative momentum k; this is because of the simplicity of the low-E Dirac eqtn., so that apart from the linear derivative ∂_μ we only have the function $A_\mu(x)$.

Notice also that we have yet to do any kind of gradient expansion; this will come when we go over to the functional integration over $A_\mu(x)$.

INTEGRATED PROPAGATOR $G_2(x,x')$ IN IR REGIME:

We now wish to integrate out the A-field, to get an expression for $G_2(x,x')$. This is the real point of deriving the result in (276); it forms a platform for deriving an expression for the propagator which does not explicitly involve a sum of graphs. Thus we will be evaluating an expression of form

$$G_2(x,x') = \int \mathcal{D}A^\mu \, \mathcal{P}[A^\mu] \, G_2(x,x'|A) \tag{277}$$

where $\mathcal{P}[A^\mu]$ is a weighting functional for different configurations of $A^\mu(x)$; essentially a probability amplitude distribution.

What do we choose for the weighting functional? Let's recall what we found in previous chapters for the generating functional & propagators in QED. The generating functional can be written as (cf chapter 3.4, eqtn. (168)):

$$Z_{EM}[J_\mu] = \int \mathcal{D}A_\mu \, e^{\frac{i}{\hbar}\left\{ \sum_q \frac{1}{2} A^\mu_q (D^0_{\mu\nu}(q))^{-1} A^\nu_{-q} - J_\mu(q) A^\mu_{k+q} \right\}} \tag{278}$$

where one assumes the existence of a fixed source $J_\mu(x)$ which we can, if we wish,

associate with the Dirac fermions via the usual relationship between $\bar{\psi}(x)$, $\psi(x)$, and the Dirac current $J_\mu(x)$.

From a comparison of (277) and (278), we might naively suppose that the weighting factor $\mathcal{P}[A^\mu]$ is then just given by

$$\mathcal{P}[A^\mu] = \mathcal{P}_0[A^\mu] = e^{\frac{i}{2\hbar}\sum_q A^\mu_q (D^0_{\mu\nu}(q))^{-1} A^\nu_{-q}} \tag{279}$$

Now actually this cannot be right in general, because we know that the photon propagator itself is renormalized by the coupling to fermions - we will return to this point below. But let's first just try this out, to see what happens. Putting (279) into (277), we have

$$\mathcal{G}_2(x,x') = \int \mathcal{D}A^\mu \, e^{\frac{i}{2\hbar}\sum_q A^\mu_q (D^0_{\mu\nu}(q))^{-1} A^\nu_{-q}} \; \mathcal{G}_2(x,x'|A^\mu_q) \tag{280}$$

or, if we want to be more sophisticated, the renormalized photon propagator $D_{\mu\nu}(q)$ in place of $D^0_{\mu\nu}(q)$ in (280).

Notice that we have no Fadeev-Popov determinant in (280); again, I will return to this point below.

Let's now see what we get from this. The Gaussian functional integral is trivial; writing:

$$\begin{aligned} \mathcal{F}_S(p,q|A) &= f_S^\mu(p,q)\, A_\mu(q) \\ &= \left(e\int_0^S ds'\, e^{is'p^\nu q_\nu} p^\mu\right) A_\mu(q) \end{aligned} \tag{281}$$

we have

$$\begin{aligned} \mathcal{G}_2(p) &= i\int_0^\infty ds \int \mathcal{D}A^\mu_q \; e^{\frac{i}{2}\sum_q A^\mu_q (D^0_{\mu\nu}(q))^{-1} A^\nu(-q) + i f_S^\mu(p,q) A_\mu(q)} \; e^{isG_0^{-1}(p)} \\ &= i\mathcal{N}\int_0^\infty ds\; e^{isG_0^{-1}(p)} \; e^{-\frac{i}{2}\sum_q f_S^\mu(p,q) D^0_{\mu\nu}(q) f_S^\nu(p,-q)}. \end{aligned} \tag{282}$$

where the normalization factor $\mathcal{N}$ will drop out later, so we henceforth ignore it. We can write the interaction term in (282) in various ways, viz.,

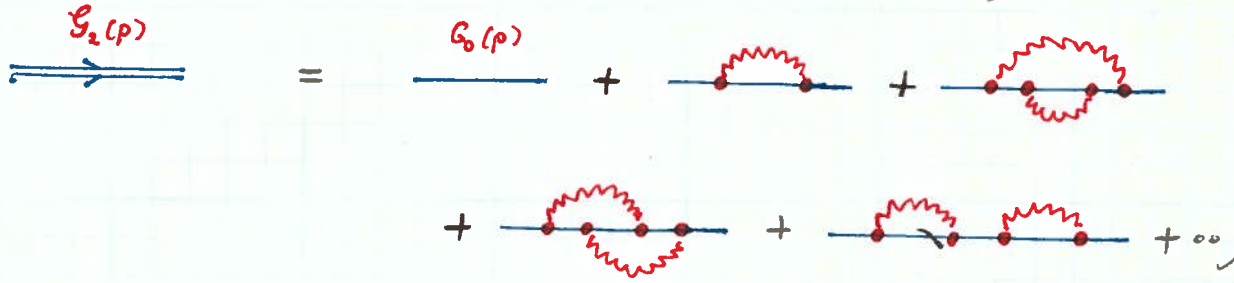
$$\begin{aligned} \chi_S(p) &= \frac{1}{2}\sum_q f_S^\mu(p,q) D^0_{\mu\nu}(q) f_S^\nu(p,-q) \\ &= \frac{e^2}{2}\sum_q \int_0^S ds_1 \int_0^S ds_2 \; e^{i(s_1-s_2)p^\mu q_\mu} \; p^\mu D^0_{\mu\nu}(q) p^\nu \\ &\equiv \frac{e^2}{4}\sum_q \int_0^S ds_1 \int_0^{s_1} ds_2 \; e^{i(s_1-s_2)p^\mu q_\mu} \; p^\mu D^0_{\mu\nu}(q) p^\nu \\ &= e^2 \sum_q \frac{1-e^{isp^\mu q_\mu}}{(p^\mu q_\mu)^2} \; p^\mu D^0_{\mu\nu}(q) p^\nu \end{aligned} \tag{283}$$

and we will use the Ikot form, to write

$$G_2(p) = i\int_0^\infty ds\, e^{isG_0^{-1}(p)}\, e^{-i\chi_s(p)}$$

$$= i\int_0^\infty ds\, \exp\left\{ i\left[(\gamma^\mu p_\mu - m + i\delta) - e^2 \sum_q \frac{1 - e^{isp^\mu q_\mu}}{(p^\mu q_\mu)^2}\, p^\mu D^0_{\mu\nu}(q)\, p^\nu \right]\right\} \qquad (284)$$

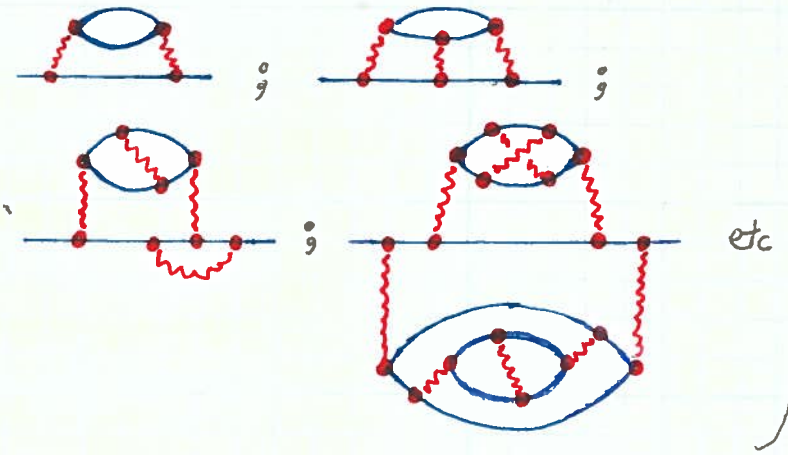
Now let's stop and look at this result, since it is actually quite interesting. The phase factor $\chi_s(p)$ is the sort of thing one sees often in calculations of wave propagation through a medium - we do not pause here to explore this - and from the point of view of perturbation theory it has the effect of summing over an infinite set of graphs, as we see in the figure.



(285)

This graphical sum is easily demonstrated starting from the 3rd form for $\chi_s(p)$ in (283). We notice that, since all the internal fermion lines are bare fermion lines, we sum over both reducible and irreducible graphs in (285). The graphical sum also makes clear what was noted above, viz., that we are only using free photons in the calculation. Thus all the following graphs, shown in the figure, are missing from the result:

MISSING GRAPHS:



(286)

We are also missing all the disconnected graphs, but these are cancelled by the normalization factor anyway.

Thus we see that we have a certain approximation here for the fermion propagator, in which we have an arbitrary number of photon loops, but no fermion loops at all.

Why is this result not exact? The reason can be seen if we compare with what we would do if we wanted to get an exact result from this

tactic of integrating over all configurations of $A^\mu(q)$. This would involve the kind of calculation we did in chapter 8.3, where we coupled 2 scalar fields, and then functional integrated over one of them while freezing the other. Now if we did this, we would get the same result as above for $p^2 \ll m^2$, but for for larger fermionic momentum the results would be more complicated, because then higher momentum terms would enter into the results. To see this in this WKB formalism, note that the Fourier-transformed function $\mathcal{G}_2(p,x|A)$ satisfies

$$[\gamma^\mu(i\partial_\mu - p_\mu + eA_\mu) - m]\,\mathcal{G}_p(x|A) = -1 \qquad (287)$$

and to get the result for $\mathcal{I}_s$ given above we have assumed that $\gamma^\mu p_\mu - m$ is small. If we drop this assumption, then the results are more messy, and indeed we find an infinite hierarchy of eqtns again for $\mathcal{I}_s$, when expanded in powers of e.

Physically one can understand what is going on here in a fairly simple way. When the photon momenta are small, an electron moving through a vacuum cannot couple to virtual electron-positron pairs, because these require an energy $\geq 2mc^2$ to create. Thus none of the graphs with the internal fermion loops can play a role, and all that is left is the set of graphs shown in (285). However once the threshold energy at $2mc^2$ is reached, the imaginary part of the fermion self-energy will start to be strongly affected by this process.

We also notice that in this calculation, I have paid no attention to the whole problem of pathologies of the functional integration over $A_\mu(q)$, with the necessity for Fadeev-Popov determinants, etc. A proper calculation would start from a determination of $\mathcal{Z}$, with the appropriate introduction of the FP determinant, followed by functional differentiation, etc. But, the key is that for this calculation it doesn't matter. This is because ultimately we have to fix a gauge when we do the calculation of $\mathcal{G}_2(p)$; it is not a gauge-invariant quantity anyway. What we are actually after is the singularity structure of the propagator, which we expect to be gauge-independent because it is physical.

The next step in this discussion would be the calculation of (284), by direct integration over $\int d^4q$, followed by the integration $\int ds$ over proper time. Actually we do not do this here, since it is rather lengthy and unilluminating, but simply quote the answer:

$$\mathcal{G}_2(p) \xrightarrow[p^2 \ll m^2]{} \frac{1}{\gamma^\mu p_\mu - m + i\delta}\left|1 - \frac{\gamma^\mu p_\mu}{m}\right|^\beta \qquad (288)$$

where the parameter β is given by

$$\beta = (3-\alpha)\frac{\alpha_{QED}}{2\pi} \qquad (289)$$

where α is the gauge fixing parameter in the photon propagator, and α_{QED} is just the dimensionless coupling in QED (i.e., $\alpha \approx 1/137$).

Because α_{QED} is so small, the effect of the IR divergence on the electron propagator at low energy is rather small. However, as noted above, the effect is rather noticeable in astrophysics and in hot charged plasmas – it is also crucial to accelerator design.

A more detailed analysis reveals that what is going on here is the emission of very large numbers of soft photons in the IR limit. Notice that as $q \to 0$ in (284), the integral Σ is proportional to $\int d^4q/q^4$, giving the log divergence that leads to the power law in (283), upon exponentiation. This means that at an energy q we have a number of $1/q$ photons. This is why the classical calculation of Bremsstrahlung is exact in the low energy limit, because a bosonic field becomes classical when its modes are highly populated.

QED for a FINITE-DENSITY FERMIONIC SYSTEM : Up to now we have been looking at a single Dirac fermion coupled to photons. But a much more realistic and interesting problem involves a finite density of fermions – this can describe situations ranging from the early universe, the interior of stars (conventional, white dwarf and neutron stars) and other plasmas on earth or in space, down to conducting condensed matter systems (ie., metallic systems).

The situation now changes radically, for 2 reasons. First, we break Lorentz invariance – the Fermi sea fixes a reference frame. Second, there is no longer an energy gap for the production of particle/anti-particle pairs – as we saw in the discussion of Fermi liquids, we can excite particle-hole pairs of arbitrarily low energy.

To deal with these changes we need to both modify the photon propagator, and modify the eqtn. of motion for the fermion propagator $G_2(xx'|A)$ in the presence of some gauge configuration $A_\mu(x)$. Let's first look at the photon propagator – the fully renormalized form for this we have already seen, and is given by

$$D_{\mu\nu}(q) = D^0_{\mu\nu}(q) + D^0_{\mu\alpha}(q)\,\Pi_{\alpha\beta}(q)\,D_{\beta\nu}(q) \qquad (290)$$

where the polarization part $\Pi_{\mu\nu}(q)$ was discussed briefly in the last chapter.

In what follows I don't give any detailed discussion of $\Pi_{\mu\nu}(q)$, which for a charged liquid of fermions is a fairly lengthy task, but simply remark that the charged part of this function (ie., $\Pi_{00}(q)$, and $\Pi_{0j}(q)$, where j runs over the spatial indices) are short-range in real space, because of charge screening. However the magnetic part $\Pi_{ij}(q)$ is long-ranged – one cannot easily screen magnetic fields (there is no "Faraday cage" for magnetic fields). This means that the low-q behaviour of $\Pi_{ij}(q)$ is important, and indeed it can be shown fairly easily that it will yield the form

$$D_{ij}(q) = (\delta_{ij} - \hat{q}_i\hat{q}_j)\,\frac{|\underline{q}|}{\chi|\underline{q}| - i\gamma|\omega|} \qquad (291)$$

in the Landau gauge ($\alpha = 0$) where χ and γ are dimensionful functions which

depend on the details of the liquid of fermions involved: the densities and masses of these fermions, the spatial dimension, and so on. We note that the components $\Pi_{00}(q)$ and $\Pi_{01}(q)$ have been eliminated here (playing no role in the $q \to 0$ IR regime), and we are dealing here purely with transverse photons (ie., magnetic degrees of freedom). One can calculate χ and γ most easily in a RPA, in which $\Pi_{\mu\nu}(q) \to \Pi_0(q)$, the simple particle-hole bubble; a more sophisticated calculation changes χ and γ, but not the form of (291). Nor is it changed if one changes the spatial dimension, nor even if one substitutes heavy quarks instead of the electrons, and, eg., gluons instead of the photons. Thus the following calculation is just as relevant to QCD as it is to QED, and can be applied to high-density astrophysical phenomena, as well as to the early universe.

In what follows we discuss a simple example of the kind of calculation that can be done. To make the integrations easier, we reduce the number of spatial dimensions to 2; this has the benefit of making the calculation relevant to systems like 2DEG's (2-dimensional electron gases, used in electronic devices) or the FQHE liquids (FQHE = Fractional Quantum Hall Effect); the model has also been discussed for strongly-correlated systems in 2 dimensions such as high-T_c superconductors.

The effective Hamiltonian we are interested in for the system is then

$$\mathcal{H}_{\text{eff}} = \int d^D r \left\{ \frac{1}{2m} \psi^\dagger(x) \left[(i\underline{\nabla} - g_0 \underline{A}(x))^2 - \mu \right] \psi(x) + \left[(\partial_t \underline{A})^2 + (\nabla \times \underline{A})^2 \right] \right\} \tag{292}$$

where all electric field effects have been eliminated, so we deal with $\underline{A}(x)$. Then the 1-particle fermion propagator satisfies the eqtn.

$$\left[i\partial_t - \frac{1}{2m} (-i\underline{\nabla} - g_0 \underline{A}(x))^2 + \mu \right] \mathcal{G}_2(xx'|\underline{A}) = -\delta(x-x') \tag{293}$$

so that its partial Fourier transform over the relative variable $x-x'$, defined by

$$\left.\begin{aligned} \mathcal{G}_2(xx'|\underline{A}) &= \sum_{\underline{k}} \int \frac{d\epsilon}{2\pi} \mathcal{G}(\underline{k}, \epsilon; x|\underline{A})\, e^{i[\underline{k}\cdot(\underline{x}-\underline{x}') - \epsilon(t-t')]} \\ &\equiv \sum_k \mathcal{G}_k(x|\underline{A})\, e^{ik(x-x')} \end{aligned}\right\} \tag{294}$$

which satisfies (again defining $p = \hbar k$):

$$\left[i\partial_t - \epsilon - \epsilon_p^0 - \tfrac{p}{m}(-i\underline{\nabla} - g_0 \underline{A}) - \frac{1}{2m}(-i\underline{\nabla} - g_0 \underline{A})^2 \right] \mathcal{G}_k(x|\underline{A}) = -1 \tag{295}$$

Then we go through the same manoeuvres as before - writing

$$\mathcal{G}_k(x|A) = i \int_0^\infty ds\, \mathcal{G}_k(x;s|A) \tag{296}$$

we write

$$\mathcal{G}_k(x;s|A) = e^{i[s G_0^{-1}(k) + \mathcal{V}_s(k, x|\underline{A})]} \tag{297}$$

and expand Ψ in powers of the bare coupling g_0, as

$$\Psi_s(k,x|\underline{A}) = \sum_n g_0^n \psi_n(s;kx|\underline{A}) \tag{298}$$

Then, upon substitution, we find that

$$\psi_n = \int_0^s ds' e^{-i(s-s')\left[i\partial_t + \frac{\nabla^2}{2m} + i\frac{k_0\cdot\nabla}{m}\right]} \sum_{r=1}^{n-1}\left[\underline{\nabla}\psi_r(s)\cdot\underline{\nabla}\psi_{n-r}(s') - 2\underline{A}\cdot\underline{\nabla}\psi_{n-1}(s)\right] - \underline{A}^2\delta_{n2} \tag{299}$$

with the lowest term given by

$$\psi_1(s;kx|\underline{A}) = \frac{g_0}{m}\sum_q e^{i(\underline{q}\cdot\underline{x}-\omega t)}\frac{1-e^{-is(\omega-\Omega_{\underline{p}}(\underline{q}))}}{\omega-\Omega_{\underline{p}}(\underline{q})}\underline{p}\cdot\underline{A}(x) \tag{300}$$

where we define:

$$\left.\begin{aligned}\Omega_{\underline{p}}(\underline{q}) &= \epsilon^0_{\underline{p}+\underline{q}} - \epsilon^0_{\underline{p}} = \frac{1}{m}\left(\underline{p}\cdot\underline{q} + \tfrac{1}{2}|\underline{q}|^2\right)\\ &\sim \underline{p}\cdot\underline{q}/m\end{aligned}\right\} \tag{301}$$

and drop the $|\underline{q}|^2$ term because we assume q^2 is small — again, we work in the IR regime.

Again, to recover a result for $\mathcal{G}_2(p) = \mathcal{G}_2(p,\epsilon)$ we functionally integrate over the gauge field $\underline{A}(x)$; but this time we use a weighting factor

$$\mathcal{P}[A^\mu] = e^{i/2h\sum_q A^i(q)D^{-1}_{ij}(q)A^j(-q)} \tag{302}$$

where $D_{ij}(q)$ is the renormalized transverse propagator in (291). Carrying out the Gaussian integration as before, and using the lowest approximation ψ_1 in (300) for $\Psi_s(k,x|\underline{A})$, we find then

$$\mathcal{G}_2(\underline{p},\epsilon) = i\int_0^\infty ds\, e^{-is\,G_0^{-1}(\underline{p},\epsilon)}\exp\left\{i\frac{g_0^2}{m^2}\sum_{\underline{q}}\int\frac{d\omega}{2\pi}D_{ij}(q)p_ip_j\,\frac{1-e^{-is(\omega-\Omega_{\underline{p}}(\underline{q}))}}{(\omega-\Omega_{\underline{p}}(\underline{q}))^2}\right\} \tag{303}$$

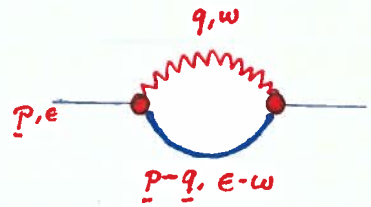
which we can compare with (284).

Although (303) looks very similar to (284), the subsequent development proceeds rather differently because the propagator $D_{ij}(q)$ is very different from the bare photon propagator.

2ND-ORDER PERTURBATION THEORY: Before doing the full phase integral in (303) (for that is what it is), it is very useful to get a quick picture of what happens in the lowest-order diagram in our theory. Thus we consider the self-energy contribution

$$\Delta\Sigma'_{\underline{p}}(\epsilon) = \frac{g_0^2}{m^2}\sum_{\underline{q}}\int\frac{d\omega}{2\pi}D_{ij}(\underline{q},\omega)\left(p_i+q_i/2\right)\left(p_j+q_j/2\right)G_0(\underline{p}-\underline{q},\epsilon-\omega) \tag{304}$$

which can be derived from standard diagrammatic rules for the diagram shown in the figure (see Appendix on diagrams).



$$= \Delta\Sigma_{\underline{p}}(\epsilon) \qquad (305)$$

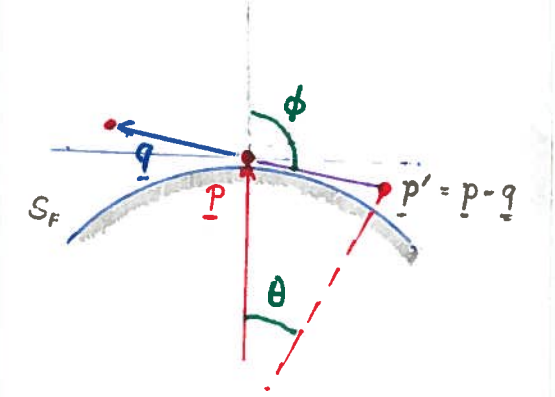
The finite-T generalization of this result, for a complex frequency z, is given in the usual way by

$$\Delta\Sigma_{\underline{p}}(z) = \frac{g_0^2}{N^2} \sum_{\underline{q}} \int \frac{dw}{2\pi} D_{ij}(\underline{q},w)\,(p_i + q_i/2)(p_j + q_j/2) \left[\frac{1 - f_{\underline{p}-\underline{q}} + n(w)}{z - \epsilon^0_{\underline{p}-\underline{q}} - w} + \frac{f_{\underline{p}-\underline{q}} + n(w)}{z - \epsilon^0_{\underline{p}-\underline{q}} + w} \right] \qquad (306)$$

We can now make a number of simplifications here. Notice first that

$$(\delta_{ij} - \hat{q}_i \hat{q}_j)\left[(p_i + q_i/2)(p_j + q_j/2)\right] = |\underline{p} \times \hat{\underline{q}}|^2 = p^2 \sin^2\phi \qquad (307)$$

where the geometry of the fermion scattering from $\underline{p}$ to $\underline{p}-\underline{q}$ is shown in the figure - the Fermi surface S_F is shown, along with its mirror reflection across the line running through the point $\underline{p}$ in momentum space, perpendicular to the vector $\underline{p}$, measured from the centre of the Fermi sphere. Since we are working in 2-d, the Fermi sphere becomes a circle.



GEOMETRY OF FERMION SCATTERING IN $\Delta\Sigma_{\underline{p}}(\epsilon)$

It is assumed in this calculation that the fermion with momentum $\underline{p}$ is above the Fermi energy μ; the intermediate state with momentum $\underline{p}-\underline{q}$ must also be. If the particle with momentum $\underline{p}$ is close to the Fermi surface S_F (and we are interested in precisely this case), then we see that the angle ϕ is roughly $90°$, and therefore, defining

$$\left.\begin{aligned} q_{\parallel} &= q\cos\phi \\ q_{\perp} &= q\sin\phi \end{aligned}\right\} \qquad (308)$$

we see that

$$\left.\begin{aligned} |\underline{q}| \sim q_{\perp} &\sim 2p_F \sin\theta/2 \sim p_F\theta \\ \sin^2\phi &\sim 1 \\ \Omega_{\underline{p}}(\underline{q}) &\sim \frac{\underline{p}\cdot\underline{q}}{m} = v_F q_{\parallel} \end{aligned}\right\} \qquad (309)$$

Now we don't actually want to do an exact calculation of $\Delta\Sigma_{\underline{p}}(\epsilon)$ in (306), just an approximate one, in order to see what sort of answer we get, and to help us see how to exclude the more complicated integrals in (303). To do this approximate evaluation of $\Delta\Sigma_{\underline{p}}(\epsilon)$ (or indeed an exact one) it is easier to first

find $\mathrm{Im}\,\Sigma_p(\epsilon)$, and then deduce the behaviour of $\mathrm{Re}\,\Sigma_p(\epsilon)$ from this. It then helps to write $\mathcal{D}_{ij}(q,\omega)$ in the form

$$\left.\begin{aligned}
\mathcal{D}_{ij}(\underline{q},\omega) &= (\delta_{ij}-\hat{q}_i\hat{q}_j)\,\mathcal{D}_\perp(\underline{q},\omega) \equiv (\delta_{ij}-\hat{q}_i\hat{q}_j)\,\frac{|\underline{q}|}{\chi|\underline{q}|-i\gamma|\omega|} \\
\mathcal{D}_\perp(\underline{q},\omega) &= \int_{-\infty}^{\infty}\frac{d\Omega}{\pi}\,\mathrm{Im}\,\mathcal{D}_\perp(\underline{q},\Omega)\,\frac{1}{\Omega-(\omega+i\delta)} \\
&\equiv \int_0^\infty \frac{d\Omega}{\pi}\,\mathrm{Im}\,\mathcal{D}_\perp(\underline{q},\Omega)\left[\frac{1}{\Omega-\omega-i\delta}-\frac{1}{\Omega+\omega+i\delta}\right]
\end{aligned}\right\} \quad (310)$$

where the spectral function of $\mathcal{D}_\perp(\underline{q},\omega)$ is, from (309):

$$\mathrm{Im}\,\mathcal{D}_\perp(\underline{q},\Omega) \simeq \frac{-\gamma\Omega q_\perp}{\chi^2 q_\perp^6+\gamma^2\Omega^2} \quad (311)$$

It then follows that we can write on the mass shell (i.e., for $\epsilon\to\epsilon_p^\circ$), that

$$\left.\begin{aligned}
\mathrm{Im}\,\Delta\Sigma_{\underline{p}}(\epsilon_p) &= -\pi\left(\frac{g_0 p_F}{m}\right)^2 \sum_{\underline{q}} \sin^2\phi \int_0^\infty \frac{d\Omega}{\pi}\,\mathrm{Im}\,\mathcal{D}_\perp(\underline{q},\Omega)\Big[\big(1-f_{\underline{p}-\underline{q}}+n(\Omega)\big)\,\delta\big(\Omega_{\underline{p}}(\underline{q})-\Omega\big) \\
&\qquad + \big(f_{\underline{p}-\underline{q}}+n(\Omega)\big)\,\delta\big(\Omega_{\underline{p}}(\underline{q})+\Omega\big)\Big] \\
&\xrightarrow{|\underline{p}|\gtrsim p_F} -\pi\left(\frac{g_0 p_F}{m}\right)^2 \sum_{\underline{q}} \int_0^\infty \frac{d\Omega}{\pi}\,\mathrm{Im}\,\mathcal{D}_\perp(\underline{q},\Omega)\,\big(1-f_{\underline{p}-\underline{q}}+n(\Omega)\big)\,\delta(v_F q_\parallel-\Omega)
\end{aligned}\right\} \quad (312)$$

where the latter form drops the term $\propto \delta(\Omega_{\underline{p}}(\underline{q})+\Omega)$, which only plays a role when $|\underline{p}|<p_F$. Let's now go to the limit $kT\to 0$, and also simplify the momentum integration by writing

$$\sum_{\underline{q}} = \frac{1}{(2\pi)^2}\int q\,dq\int d\phi = \frac{N(0)}{2\pi p_F}\int dq_\perp \int d\epsilon' \quad (313)$$

with $\epsilon' \equiv \epsilon^\circ_{\underline{p}-\underline{q}} \simeq p_F/m\,(p'-p_F)$, and where the Fermi surface density of states is given in 2-dimensions by $N(0) = m/2\pi$. Because $kT=0$, we can then write

$$\left.\begin{aligned}
\int_{-\infty}^{\infty} d\epsilon'\,\big(1-f(\epsilon')+n(\Omega)\big)\,\delta(\epsilon-\epsilon'-\Omega) &= f(\Omega-\epsilon)+n(\Omega) \\
&\xrightarrow{kT=0} \theta(\Omega)\,\theta(\epsilon-\Omega)
\end{aligned}\right\} \quad (314)$$

so we can now simplify (312) to

$$\mathrm{Im}\,\Delta\Sigma_{\underline{p}}(\epsilon_p) \simeq \left(\frac{g_0}{m}\right)^2 \frac{N(0)\,p_F}{2\pi}\int_0^{\epsilon_p} d\Omega \int_0^\infty dq_\perp\,\mathrm{Im}\,\mathcal{D}_\perp(q_\perp,\Omega) \quad (315)$$

which we can then calculate using (311). The integrals are easily done, and we get the result

$$\mathrm{Im}\,\Delta\Sigma_p(\epsilon_p) \sim \tilde{g}^2 |\epsilon_p|^{2/3} \tag{316}$$

with the renormalized coupling $\tilde{g}$ given by

$$\tilde{g}^2 = g_0^2 \, \bar{C} \, \frac{p_F}{m \chi^{2/3} \gamma^{1/3}} \tag{317}$$

with $\bar{C} = \int_0^\infty dx \, (1+x^6)^{-1}$.

Using dispersion relations we can then deduce the form of $\mathrm{Re}\,\Delta\Sigma_p(\epsilon)$. The conclusion of this analysis is that

$$\left.\begin{aligned} \mathrm{Im}\,\Delta\Sigma_p(\epsilon) &\sim \tilde{g}^2 |\epsilon|^{2/3} \\ \mathrm{Re}\,\Delta\Sigma_p(\epsilon) &\sim \tilde{g}^2 |\epsilon|^{2/3} \end{aligned}\right\} \quad \Delta\Sigma_p(\epsilon) \sim e^{i\phi} \tilde{g}^2 |\epsilon|^{2/3} \tag{318}$$

where $\tan\phi = \mathrm{Im}\,\Delta\Sigma_p(\epsilon) / \mathrm{Re}\,\Delta\Sigma_p(\epsilon)$.

What does this result tell us? Several things. First, the result completely evades all the arguments that we went through in chapter 4, which indicated that the effect of interactions on fermions near the Fermi surface was simply to renormalize things like the fermion mass ($m \to m^*$), with a quasiparticle decay rate $\sim \mathrm{Im}\,\Sigma_p(\epsilon) \sim O(\epsilon^2)$, so that the quasiparticles were well-defined as $\epsilon \to 0$, ie., near the Fermi surface. Here we have $|\mathrm{Im}\,\Delta\Sigma_p(\epsilon)| > |\epsilon|$, ie., the decay rate becomes greater than the quasiparticle energy, and the quasiparticles are no longer well defined.

However, one can immediately ask whether this result is credible - after all, if the lowest-order graph contributing to $\mathcal{G}_2(p)$ is infra-red divergent, surely we need to go to higher order? Indeed from (318) all we can infer is that

$$\mathcal{G}_2(p=p_F, \epsilon) \sim \frac{1}{\epsilon}\left[1 - e^{i\phi} \tilde{g}^2 |\epsilon|^{-1/3} + O(\tilde{g}^4) \ldots\right] \tag{319}$$

and from our experience so far, we can expect that the higher-order graphs will be even more singular. Notice that in 4-d QED, the higher-order contributions went like $g^{2n}(\ln\epsilon)^n$; here it looks like they are going to go like $\tilde{g}^{2n}|\epsilon|^{-n/3}$, a much worse power law divergence - in fact, we will see in the next chapter that such contributions are rather dangerous to the theory.

Thus we now really need a non-perturbative treatment, and this is what (303) provides us with.

EVALUATION OF FULL PHASE INTEGRAL: If we look at (303), we see immediately that it has the form of a "singular phase integral", ie., an integral over a phase (a function the proper time s) which is divergent, because of the energy denominators in the exponent. We should not be too bothered by this - after all, if we consider the simple path integral for an electron moving in the field of an attractive proton, the propagator is

$$\mathcal{G}_{\text{Coulomb}}(2,1) = \int_1^2 \mathcal{D}\underline{r}(t) \, e^{i/\hbar \int_{t_1}^{t_2} dt \left[\frac{1}{2} m \dot{\underline{r}}^2 + e^2/|\underline{r}|\right]} \tag{320}$$

and the phase $\frac{1}{\hbar} S[\underline{r}]$ is definitely divergent for paths which approach the origin!

But we don't worry about this because we know that such paths make little contribution to the path integral – they involve extremely fast oscillations (corresponding to the diverging kinetic energy in this regime) which average to zero. Moreover, we would never think of expanding the exponential as a power series, ie., writing

$$G_{Coulomb}(2,1) = \int_1^2 \mathcal{D}r(t)\, e^{\frac{i}{\hbar}\int_1^2 dt\, \frac{m}{2}\dot{\underline{r}}^2(t)} \left[1 + \int_{t_1}^{t_2} dt\, \frac{e^2}{|\underline{r}(t)|} + \frac{1}{2}\int_1^2 dt \int_1^2 dt'\, \frac{e^4}{|\underline{r}(t)||\underline{r}(t')|} + \text{etc} \right] \tag{321}$$

unless we were feeling silly – every term is formally divergent, and so the series makes no sense. Thus the integral (320) is well-defined, but not the series expansion (321).

In the same way the series in (319) should not be taken seriously, but merely as a signal that we need to go back to the full phase integral – that perturbation theory is not well-behaved and not meaningful for this problem.

From our evaluation of the momentum integrals for the 2nd-order perturbative self-energy $\Delta\Sigma_p(\epsilon)$, we see that we can write the phase $\varphi_s(p)$ in the formula (303) for $\mathcal{G}_2(p)$, viz.,

$$\mathcal{G}_2(p) = i\int_0^\infty ds\, e^{-is\mathcal{G}_0^{-1}(p)}\, e^{i\varphi_s(p)} \tag{322}$$

in the form

$$\left.\begin{aligned} \varphi_s(p) &= \frac{g_0^2}{m^2}\sum_{\underline{q}}\int\frac{dw}{2\pi}\, \mathcal{D}_{ij}(\underline{q},w)\, p_i p_j\, \frac{1-e^{-is(w-\Omega_{\underline{p}}(\underline{q}))}}{(w-\Omega_{\underline{p}}(\underline{q})+i\delta)^2} \\ &= \frac{1}{(2\pi)^2}\left(\frac{g_0 p_F}{m}\right)^2 \int\frac{dw}{2\pi}\int_0^\infty dq_\parallel \int_0^\infty dq_\perp\, \frac{1-e^{-is(w-q_\parallel v_F)}}{(w-q_\parallel v_F+i\delta)^2}\, \mathcal{D}_\perp(q_\perp,w) \end{aligned}\right\} \tag{323}$$

so that the momentum integrals factorize into separate integrals over $q_\parallel$ and $q_\perp$. If we carry out these two integrals, we find that

$$\varphi_s(p) = \frac{1}{(2\pi)^2 v_F}\,\frac{1}{(\chi^2\gamma)^{1/3}}\,\frac{e^{-i\pi/6}}{3^{3/2}}\left(\frac{g_0 p_F}{m}\right)^2 \int_0^\infty \frac{dw}{w^{1/3}}\left[\frac{1-e^{isw}}{w} - is\left(Ci(sw)+iSi(sw)\right)\right]$$

$$= \bar{g}^2 (is)^{1/3} \tag{324}$$

with the renormalized coupling

$$\bar{g}^2 = g_0^2\left[\frac{\Gamma(2/3)}{8\pi^2\sqrt{3}}\,\frac{1}{(\chi^2\gamma)^{1/3}}\,\frac{p_F}{m}\right] \tag{325}$$

Thus we finally have a proper time integral to do; substituting (324) into (322),

we get

$$\left.\begin{aligned} \mathcal{G}_2(p_F,\epsilon) &= i\int_0^\infty ds\, e^{is(\epsilon+i\delta)}\, e^{-\tilde{g}^2(is)^{1/3}} \\ &= -i\frac{\pi}{\epsilon}\left[\frac{d^2}{dz^2}\,Hi(z)\right]\Big|_{z=e^{4\pi i/3}\tilde{g}^2\epsilon^{-1/3}} \end{aligned}\right\} \quad (326)$$

where the function $Hi(z)$ is defined in terms of the Airy functions $Ai(z)$ and $Bi(z)$ as

$$\left.\begin{aligned} Hi(z) &= \frac{1}{\pi}\int_0^\infty dx\, e^{-(x^3/3 - xz)} \\ &= \frac{2}{3}Bi(z) + \int_0^z dx\,[Ai(x)Bi(z) - Ai(z)Bi(x)] \end{aligned}\right\} \quad (327)$$

This result shows that the renormalized photons in the system have a drastic effect on the fermion propagator near the Fermi surface. The Airy functions have an essential singularity $\sim \exp\{\frac{2}{3}z^{3/2}\}$ as $z\to\infty$, but a perturbative expansion as $z\to 0$. However, as we see from (326), the limit $z\to\infty$ corresponds to $\epsilon\to 0$, and so we find that $\mathcal{G}_2(p_F,\epsilon)$ has an essential singularity as $\epsilon\to 0$, but a perturbative expansion at high energy. From (326) we find

$$\mathcal{G}_2(p_F,\epsilon) = \left\{\begin{array}{ll} -\pi^{-1/2}\,\tilde{g}^{3/2}\,\frac{1}{\epsilon^{5/4}}\,e^{-\frac{2}{3}\tilde{g}^3/\epsilon^{1/2}} & (\text{low } \epsilon:\ \epsilon \ll \tilde{g}^6) \\ -\frac{i\pi}{\epsilon}\,c_1\left[1 + c_2\tilde{g}^2\frac{i^{4/3}}{\epsilon^{1/3}} + \cdots\right] & (\text{high } \epsilon:\ \epsilon \gg \tilde{g}^6) \end{array}\right. \quad (328)$$

From these results we gain a better understanding of both the perturbative structure of the theory and of its essentially non-perturbative behaviour in the low-energy limit. From (328) we see that the structure of the high-energy series is what we would have guessed from the perturbative calculation (cf. (318) and (319)), except that this perturbative result was derived in the low-energy limit. Thus the perturbative result is very misleading. The true low-energy behaviour is far more radical — no summation can ever produce a truly singular result like the low-ϵ limit in (328). Thus this behaviour is truly non-perturbative, in the full sense of the term, viz. inaccessible by any sum over diagrams, infinite or not.

Firstly, let us note that we could have found the form (326) for $\mathcal{G}_2(p_F,\epsilon)$, within dimensionless constants, purely by dimensional analysis of the integrals over q and ω in (323). Thus, eg., we can write

$$\int_0^\infty dq_\parallel\,\frac{1-e^{-is(\omega - q_\parallel v_F)}}{(\omega - q_\parallel v_F)^2} \sim s\int_{-s\omega}^{\infty}\frac{du}{u^2}\,(1-e^{iu}) \sim sf(s\omega) \quad (329)$$

where $f(x)$ is some function; it then follows that $\varphi_s(p) \sim s\int_0^\infty \frac{d\omega}{\omega^{1/3}} f(s\omega) \sim s^{1/3}$. If we take account of the phases in (329) we get $(is)^{1/3}$, and the rest is just a renormalization of g_0^2 to $\tilde{g}^2$ (which we can only get by doing the integrals properly).

B.6.3. : THE Q.F.T. VACUUM

In ordinary quantum mechanics the vacuum state is usually held to have simple properties (although this not so obvious if we deal with a finite number of anyons). On the other hand for a system with a macroscopic number of particles, or a quantum field theory, the vacuum state becomes exceedingly complex. A number of key issues and questions arise, of which the following will be looked at in what follows:

(i) One can have spontaneous symmetry breaking of the field, ie., the symmetry of the vacuum and the symmetries underlying the excited states are different from those in the action or Lagrangian that one starts with. Often the resulting physical states are then utterly different. We have already met the results of this symmetry breaking in previous material (notably in the previous sub-section discussing ferromagnetic and anti-ferromagnetic systems); here we will discuss it in a more general way.

(ii) Quite generally, in a QFT, the vacuum energy is found in a naive calculation to be infinite, in the absence of any UV cut-off, and rather large even with one. In condensed matter systems (including stars and neutron stars) this energy is large, finite, and physically meaningful. In a relativistic QFT in flat spacetime, we subtract it off using renormalization procedure (discussed more in Ch. 7). In a curved spacetime it is not even clear how to calculate it, and various anomalies (notably the trace anomaly, discussed in the next section) are associated with it. However in all cases we can look at differences or changes in it, brought about by constraining or varying the vacuum by changing the boundary conditions on it - this leads to a physical effect known as the Casimir effect, of great theoretical and experimental interest.

(iii) One may have a multitude of different vacua in a field theory; if there is some way of passing between these, via a tunneling process, then these are not the true vacua, even if they may be topologically different, and the true vacuum will be some superposition of these. Boundary conditions are crucial here, and the topological properties of the different metastable (or stable) vacua are often (but not always) associated with a variety of interesting topological excitations like solitons. The existence of solitonic excitations in any non-linear field theory, classical or quantum, is fairly general; these can be both topological & non-topological, and they of course do not exist in any linearized theory.

(iv) Since all Lagrangians and actions are really effective ones, defined over some kind of restricted Hilbert space or restricted set of degrees of freedom, then so are their vacua. Thus, in some given theory, a change in these different restrictions (brought about by, eg., renormalization) will change the vacuum, and the excitations above it. This of course is a

limitation of our description of the system; however, it often corresponds to a real physical feature of the system, which is that for any realistic timescale, the system behaves as though it is described by some effective theory, with what is for all practical purposes the vacuum of that theory. For example, a quantum liquid behaves as if it has one vacuum, but if we change the temperature and it solidifies, it behaves as if it has a completely different vacuum. What then is the true vacuum? Even more interesting and subtle questions arise if the system is disordered. We will address some of them here, and some of them in Ch. 7.

B.6.3(a) SIMPLE VACUA & VACUUM ENERGY

Let's first go over some of the simple things we know. By "simple vacua", we mean here vacua where there is no spontaneous symmetry breaking, nor topological sectors, and of course that the background spacetime is flat. What we wish to know is how to understand the vacuum energy and its associated with vacuum fluctuations, as well as the low-energy excitations above it.

We will also assume here (and that is part of what is meant when we say "simple" vacua) that the true vacuum of the system really is what is being described by our model Lagrangian or Hamiltonian. In other words, we assume that a change of energy scale or any range in our description does not change the vacuum, so that our effective theory really is the true theory of the system. As we shall discuss, this is rarely true, but it is a useful idealization.

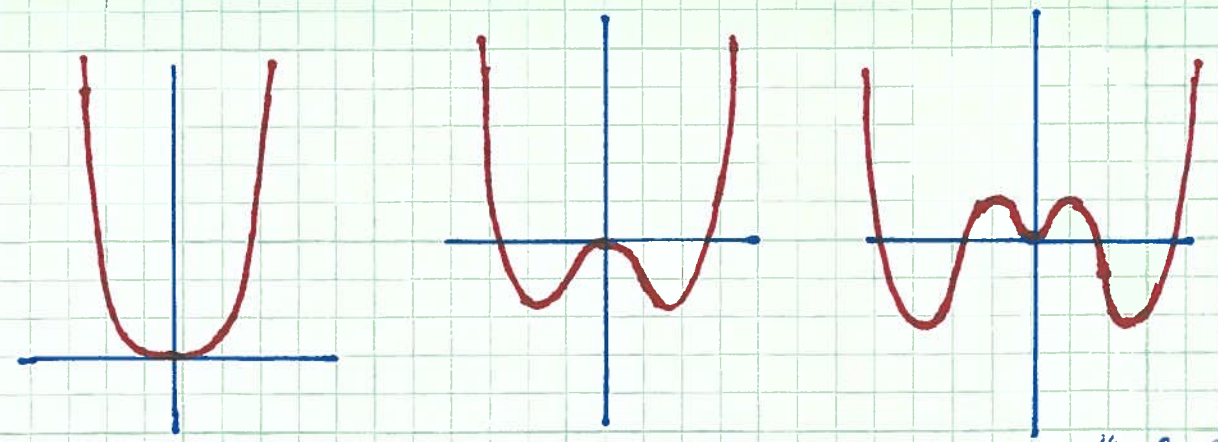
(i) SURVEY OF SIMPLE MODEL VACUA: The best way to approach this is by considering a few examples. These tell us some of the pitfalls that lie in wait for a too naive approach. We can consider both bosonic and fermionic fields.

- **SCALAR FIELD THEORY**: We consider the usual model of a massive scalar field in D dimensions, with a self-coupling term. For the sake of argument, we write the action as

$$S[\phi] = S_0[\phi] - \int d^D x \left[\frac{g_4}{4!}\phi^4(x) + \frac{g_6}{6!}\phi^6(x) \right] \qquad (1)$$

where $S_0[\phi]$ is just the standard non-interacting scalar field action, with a mass m. Note that we only include even powers of the field in the interaction; and odd power would be unbounded from below, and so the vacuum would not be stable.

It is useful to plot the form of the effective potential in (1) as a function of $|\phi(x)|$ for the cases where the energy is bounded from below.



$V(\phi) = \frac{g_4}{4!}\phi^4 + \frac{g_6}{6!}\phi^6$ $\quad$ $V(\phi) = \frac{g_4}{4!}\phi^4 + \frac{g_6}{6!}\phi^6$ $\quad$ $V(\phi) = g_4/4!\,\phi^4 + g_6/6!\,\phi^6$

$m^2, g_4, g_6 \geq 0$ $\quad$ $m^2 < 0$ $\quad g_4, g_6 > 0$ $\quad$ $m^2, g_6 > 0$ $\quad g_4 < 0$

The first example on the left is straightforward; the energy is obviously bounded from below, with a minimum at $|\phi(x)| = 0$. For small $|\phi|$, the excitations are well approximated by harmonic free field excitations, and we expect perturbation theory to work OK.

The 2nd figure has a minimum for $|\phi(x)| > 0$. As we will discuss later, this means that in the vacuum state, $|\phi(x)| \to \phi_0$; we have a broken symmetry. If we vary, for example, g_4 from positive to negative, or vary m^2 from positive to negative, we can cause a continuous transition from $|\phi| = 0$ to $|\phi| > 0$; this is a 2nd-order phase transition. A model like this essentially is Ginzburg-Landau model.

The 3rd figure has a minimum at $|\phi| = 0$, plus another one at $|\phi| > 0$. Now we can arrange a situation, by varying g_4, in which we pass from zero to finite $|\phi|$ in a discontinuous fashion; this is a 1st-order phase transition. It is clear that perturbation theory is useless here — this is a non-analytic feature (as a function of g_4) of the system behaviour.

– **PHONON FIELD THEORY**: This is very similar to the scalar field theory, if we deal with an isotropic & translationally invariant system. Thus the phenomenological scalar field theory just discussed can serve as a nice toy model for a liquid; the scalar field represents in this case either a displacement field for particles (in a smoothed or coarse-grained representation), or a density deviation field (as we will see below, for a superfluid system it will be used to describe phase excitations). Either way, we also need a conjugate momentum or velocity field to give a proper description (see Ch. 3).

Things are much more complicated in a solid. A full discussion here is well outside the scope of these notes; but we note the following key features of a comprehensive field theory of a quantum solid:

(a) Vacuum Energy : In contrast to the case of photons (which we will discuss below), there is a natural UV cut-off in the spectrum of phonon excitations, in either a liquid or a solid. This is because of the natural granularity of these systems - it makes no sense to define a phonon in a liquid for wavelengths shorter than the inter-particle spacing (which, since a liquid is fairly incompressible, is fairly well-defined); and for a disordered solid the same is also clearly true. In a crystalline solid there will be, depending on the lattice symmetry, quite a few phonon branches; but the same argument applies. As we already discussed in Ch. 3, the phonon spectrum will be linear at low energies, ie.,

$$\omega_q \sim c|q| \qquad (|\omega_q| \ll k_B \Theta_D) \tag{2}$$

and otherwise its shape depends on the microscopic structure and ordering; here Θ_D is the Debe temperature of the system.

If we now naively sum over the different phonon branches, labelled by indices λ, we have a vacuum energy

$$E_0^\phi = \tfrac{1}{2}\hbar \sum_{q,\lambda} \omega_{q,\lambda} \tag{3}$$

This "zero-point energy" is physically meaningful for a compressible solid or liquid - it is connected to the cohesion energy of the system, and when one tries to deform the system, by straining or compressing it, the change in the zero point energy is connected to the compressibility.

(b) Symmetry Group : As we noted in Ch. 3, the symmetry group of a liquid or amorphous solid is simple, so that is not so different from simple scalar field theory (with the addition of transverse modes); however, a crystalline solid may have a very complex point group symmetry, so that the free energy functional contains many different independent terms. There will then be many different phonon branches. As one changes external parameters, the system can make transitions from one vacuum state to another, having different symmetries, and with different phonon excitations. Thus the phenomenology is very complex.

(c) Defects, Dislocations, etc.: In any real solid the lattice symmetry can be broken locally; the result is a local excitation, which may or may not be mobile. In 3 spatial dimensions one can have either point defects, of various kinds, or line defects, or surface defects (ie, defects extending in some surface inside the system). Depending on the lattice symmetry, such objects can take various forms. The line defects go under various names depending on their structure - the most common are dislocations & disclinations. The surface defects may be planar (as in fault planes) or more complicated domain wall kinds of excitation, dividing the system into different regions or

domains, possibly with different vacua in the different domains.

These different excitations have a number of properties which are generic to many different field theories, relativistic or non-relativistic; these include:

- the defects typically constitute topologically stable excitations, which can only be removed at a system boundary
- if they are mobile, the constitute another set of excitations in the system, in addition to the phonons, having their own dispersion relation and quantum numbers. They will in general also have their own internal excitations.

Often one treats the vacuum of a system with N such topological excitations as distinct from that with $N' \neq N$ such excitations - this can make sense if it is difficult or impossible to pass between these vacua. The phonon system can give a good example of this; thus, eg., a solid full of dislocations has completely different properties from one without, and the local properties of, eg., the phonons will also be very different.

Later in this section we will take a brief look at the modifications to standard QFT required for the study of topological "soliton" excitations.

(d) Interactions between excitations: As discussed in Ch. 3, the field theory for phonons contains interactions between the phonons to high order. This means that both the phonon vacuum and the vacuum energy will be modified. If we study this perturbatively we will find diagrams like the ones shown in the figure.

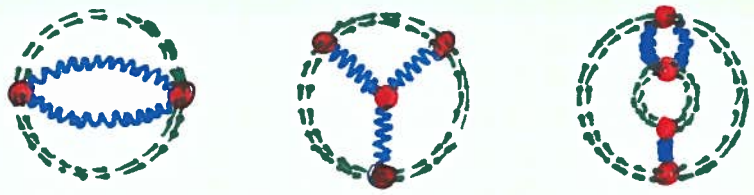
+ etc.

However, in a real solid which also possesses defects/topological excitations, we will also have interactions between the defects and the phonons. Without going into any details, this leads us to expect interaction vertices of the form shown in the figure - here the topological excitation has a propagator represented by a double hatched line.

+ etc.

We see that a perturbative treatment of the vacuum will now involve multiple contributions involving both phonon-defect and phonon-phonon interactions (and here we ignore possible direct defect-defect interactions,

which will else exist). Thus we end up with a plethora of new diagrams, as shown here we only show a small sampling.



+ etc

Thus we see that the actual vacuum energy of a solid may be very different from that given by (3); and the correlation function of the system will also be very different. The complexity of the real solid is thus quite intimidating – but this should not be a surprise, given the enormous richness to be found in studying the mechanical properties of solids.

– **NON-RELAVISTIC FERMI LIQUID**: Let us first consider the non-interacting Fermi gas. This model gives a surprisingly good description of systems ranging from Fermi gases to metals and, in relativistic form, of white dwarf stars, and dense relativistic plasmas. The vacuum energy is then given, under the assumption that the dispersion relation is (here $\sigma = \pm$ is the spin projection):

$$\epsilon^0_{k\sigma} = \hbar^2 k^2/2m \tag{4}$$

by the result

$$\left.\begin{aligned} E_0^{FG} &= \sum_{k,\sigma} \epsilon^0_{k\sigma}\, \theta(\mu - \epsilon^0_k) \\ &\equiv \frac{2}{4\pi^2}\left(\frac{2m}{\hbar^2}\right)^{3/2} \int_0^\mu d\epsilon\, \epsilon^{3/2} \\ &= \frac{1}{5\pi^2}\left(\frac{2m}{\hbar^2}\right)^{3/2} \mu^{5/2} \end{aligned}\right\} \tag{5}$$

where we have a factor 2 in the 2nd expression because of the spin sum. In the same way the number density is

$$\left.\begin{aligned} n &= \frac{2}{4\pi^2}\left(\frac{2m}{\hbar^2}\right)^{3/2} \int_0^\mu d\epsilon\, \epsilon^{1/2} \\ &= \frac{1}{3\pi^2}\left(\frac{2m}{\hbar^2}\right)^{3/2} \mu^{3/2} \end{aligned}\right\} \tag{6}$$

so that we have the well-known result that

$$E_0^{FG}/n = 3/5\,\mu \tag{7}$$

for a 3-d Fermi gas.

There is another way to think about this result - suppose we follow the common convention of putting $\mu = 0$ in these evaluations. Then the energy integral goes through as above, but with every particle energy shifted downwards by μ; we then get

$$\left.\begin{aligned} E_0^{FG}(\mu) &\rightarrow E_0^{FG}(\mu=0) \\ &= E_0^{FG}(\mu) - \mu n \\ &= -\frac{2}{15\pi^2}\left(\frac{2m}{\hbar^2}\right)^{3/2}\mu^{5/2} \end{aligned}\right\} \quad (8)$$

which now gives $$\frac{E_0^{FG}(\mu=0)}{n} = -\frac{2}{5}\mu. \quad (9)$$

The point of this exercise is to simply note that the vacuum energy is changed simply by a different choice of the zero of energy - in real systems μ is of course fixed by some external reservoir. However, we shall see that this observation is useful to remember when we come to Dirac electrons.

Consider now the effect of interactions. As we saw in Ch.5, one can write the energy functional for an interacting Fermi liquid in the form*

$$E_{FLT} = E_{FLT}^0 + \sum_\alpha \varepsilon_\alpha \delta n_\alpha + \frac{1}{2}\sum_{\alpha\alpha'} f_{\alpha\alpha'}\delta n_\alpha \delta n_{\alpha'} + \cdots \quad (10)$$

where the $\{\alpha\}$ label the quantum numbers of the system; for an isotropic and translationally invariant system, with spin-$\frac{1}{2}$ fermions, this becomes

$$E_{FLT} = E_{FLT}^0 + \sum_{\underline{p}\sigma} \varepsilon_{p\sigma}\delta n_{p\sigma} + \frac{1}{2}\sum_{\underline{p}\underline{p}'}\sum_{\sigma\sigma'} f_{pp'}^{\sigma\sigma'}\delta n_{\underline{p}\sigma}\delta n_{\underline{p}'\sigma'} \quad (11)$$

and in the absence of spin-orbit coupling, we can think of such a free energy functional as resulting from an effective interaction

$$\mathcal{H}_{int}^{eff} = \frac{1}{2}\sum_{\underline{p}\underline{p}'}\sum_{\underline{q}}\left[f_{pp'}^{s}\,\delta\hat{\rho}_{\underline{p}}(\underline{q})\,\delta\rho_{\underline{p}'}(-\underline{q}) + f_{\underline{p}\underline{p}'}^{a}\,\delta\hat{\underline{\sigma}}_{\underline{p}}(\underline{q})\cdot\delta\underline{\sigma}_{-\underline{p}'}(-\underline{q})\right] \quad (12)$$

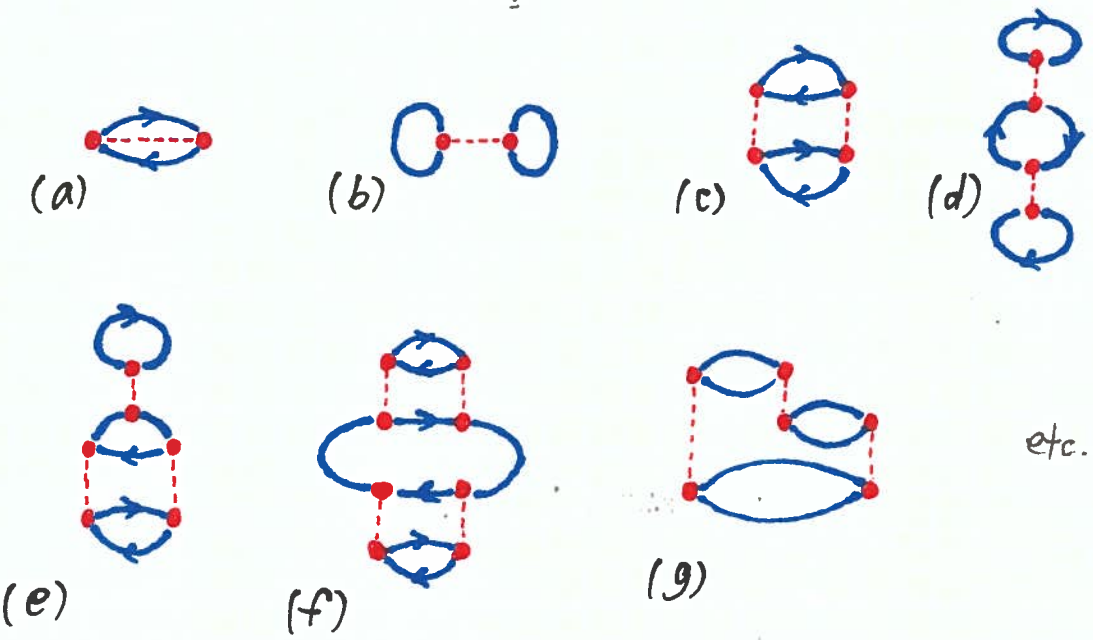
In the above eqns., as in Ch.5, we define the quasiparticle

---

* As discussed in Ch.5, this free energy functional should be written as an expansion over matrix distribution functions in the general case - we avoid this here for simplicity, assuming the matrices are diagonal.

distribution function δn_α and the quasiparticle interaction $f_{\alpha\alpha'}$, where δn_α is the deviation of the complete distribution function from its ground state unperturbed form n^0_α, ie., $\delta n_\alpha = n_\alpha - n^0_\alpha$; we define the quasiparticle energy $\varepsilon_\alpha = \delta E/\delta n_\alpha$, and the quasiparticle interaction by $f_{\alpha\alpha'} = \delta^2 E/\delta n_\alpha \delta n_{\alpha'}$; and so on. For the isotropic translationally invariant system, we have defined the density fluctuation and spin fluctuation operators $\delta\hat{\rho}$ and $\delta\underline{\sigma}$, and the spin-symmetric & spin-antisymmetric components $f^S_{pp'}$ and $f^A_{pp'}$ of $f_{pp'}$.

As I will discuss more formally below, this formulation allows us to write the vacuum energy of the system in terms of the f-function $f_{\alpha\alpha'}$, instead of just writing a perturbation expansion. The perturbation expansion just involves vacuum diagrams of the kind shown in the figure, in which some bare interaction $V(q)$ is assumed.



Each of the graphs shown in the figure has interesting features. A key point to notice here is that in some of them, all we are doing is making "self-energy insertions" into existing bare lines. Thus, eg., we can get both (e) and (f) by inserting an extra loop or bubble in interaction with the top line in diagram (c), so that we are simply renormalizing this line. In the same way (d) is produced by adding an extra loop to interact with one of the lines in (b). We will see below how to write the vacuum energy entirely in terms of renormalized propagators

In a real Fermi liquid, whether it be metallic or neutral, one also has interactions with phonons, so that the form of the theory is more complex (as are the diagrams, which we have already seen in Chapters 3 and 5). If one adds disorder, things become even more complex - one can then have interactions between the electron and defects or impurities, and things become very messy.

(ii) QUANTUM ELECTRODYNAMICS: If we ignore the QED interaction between electrons & photons, we can treat the EM and Dirac vacua separately. These 2 vacua have a few interesting features, which we quickly recall here:

(a) the bare EM vacuum in flat spacetime is just that of a set of non-interacting oscillators, with a formally divergent vacuum energy. It is very similar to a set of phonons – the difference is that one can only have transverse photon modes in a vacuum, whereas phonons always have longitudinal modes, and a natural UV cut-off.

(b) The Dirac vacuum in flat spacetime, in the conventional treatment, has a vacuum energy which is formally negative (and divergent). However, as with the Fermi gas, we can treat this as a matter of convention. In fact a more appropriate comparison is between Dirac electrons and either electrons in certain 1-d (or effectively 1-d) fermionic systems, or in higher-dimensional systems like graphene, topological insulators, or ^{3}He superfluid.

We can summarise the basic results for these 2 vacua as follows:

EM VACUUM: It is illuminating to compare the EM vacuum to the phonon vacuum, without getting too technical. Let us start with the EM field – we can get a quick result by just quantising the field $A_\mu(x)$ in the Coulomb gauge, so that $A_0(x)$ can be dropped; we then expand $\underline{A}(x)$ in photon modes, as

$$\underline{A}(x) = \sum_{\underline{q}} \frac{1}{2q_0} \sum_\lambda \underline{\varepsilon}_{\underline{q}}^\lambda \left[a_\lambda(q) e^{-iqx} + a_\lambda^\dagger(q) e^{iqx} \right] \delta(q_0 - \omega_{\underline{q}}) \qquad (13)$$

where $\sum_{\underline{q}} \equiv \int \frac{d^3q}{(2\pi\hbar)^3}$, and $\lambda = 1,2$ is the photon polarization; the $\underline{\varepsilon}_{\underline{q}}^\lambda$ obey

$$\left.\begin{aligned} \underline{q} \cdot \underline{\varepsilon}_{\underline{q}}^\lambda &= 0 \\ \underline{\varepsilon}_{\underline{q}}^{\lambda_1} \cdot \underline{\varepsilon}_{\underline{q}}^{\lambda_2} &= \delta^{\lambda_1 \lambda_2} \end{aligned}\right\} \qquad (14)$$

so that the vector $\underline{A}$ is transverse to the photon direction $\hat{\underline{q}}$. The field operators are bosonic, i.e.,

$$\left[a_{\lambda_1}(q_1), a_{\lambda_2}^\dagger(q_2) \right] = \delta_{\lambda_1,\lambda_2} \delta(\underline{q}_1 - \underline{q}_2) \qquad (15)$$

and the Hamiltonian is

$$\left.\begin{aligned} \mathcal{H}_0^{EM} &= \tfrac{1}{2}\hbar \sum_\lambda \int d^3q \, \omega_{\underline{q}} \left[a_\lambda^\dagger(q) a_\lambda(q) + a_\lambda(q) a_\lambda^\dagger(q) \right] \\ &= \hbar \sum_\lambda \int d^3q \, \omega_q \left[a_\lambda^\dagger(q) a_\lambda(q) + \tfrac{1}{2} \right] \end{aligned}\right\} \qquad (16)$$

so that at $T = 0$, when no photons are excited, we end up with a vacuum energy given by

$$E_0 = \tfrac{1}{2}\hbar \sum_\lambda \int d^3q\ \omega_{\underline{q}} = \hbar c \int d^3q\, |\underline{q}| . \qquad (17)$$

which is the famous UV catastrophe. The standard way to deal with this is to subtract off this zero-point energy, but as we will see below, it is not completely abstract (see discussion of the Casimir effect).

Let us now compare this with the phonon system. The most appropriate comparison is with the acoustic phonons, with the result given in (3). In a 3-dimensional solid, there are 3 modes ($\lambda = 0, 1, 2$), one longitudinal and 2 transverse; in an isotropic liquid, there is only one longitudinal mode. Let's consider the liquid; then we can write the phonon field in the form

$$\left.\begin{aligned} \phi(x) &= \left(\frac{\hbar c_0^2}{2\rho_0}\right)^{1/2} \sum_{\underline{q}} \frac{1}{\omega_{\underline{q}}^{1/2}} \left[b_{\underline{q}} e^{iqx} + b_{\underline{q}}^\dagger e^{-iqx} \right] \delta(q_0 - \omega_{\underline{q}}) \\ \pi(x) &= -i \left(\frac{\hbar \rho_0}{2 c_0^2}\right)^{1/2} \sum_{\underline{q}} \omega_{\underline{q}}^{1/2} \left[b_{\underline{q}} e^{iqx} - b_{\underline{q}}^\dagger e^{-iqx} \right] \delta(q_0 - \omega_{\underline{q}}) \end{aligned}\right\} \qquad (18)$$

where $\pi(x) = \partial_t \phi(x)$ is the conjugate momentum to $\phi(x)$, physically equivalent to the density deviation from the mean density.* The phonon operators obey the commutation relations

$$[b_{\underline{q}_1}, b_{\underline{q}_2}^\dagger] = \delta_{\underline{q}_1 \underline{q}_2} \qquad (19)$$

and the Hamiltonian is given by

$$\left.\begin{aligned} \mathcal{H} &= \int d^3r \left[\dot\phi(x) \frac{\partial \mathcal{L}}{\partial \dot\phi(x)} - \mathcal{L} \right] \\ &= \tfrac{1}{2} \int d^3r \left[\pi^2(x) + c_0^2 (\underline{\nabla}\phi(x))^2 \right] \\ &= \hbar \sum_{\underline{q}} \omega_{\underline{q}} \left[b_{\underline{q}}^\dagger b_{\underline{q}} + \tfrac{1}{2} \right] \end{aligned}\right\} \qquad (20)$$

where $\omega_{\underline{q}} = c_0 |\underline{q}|$, and c_0 is the sound velocity for acoustic longitudinal phonons.

We see from (20) that for longitudinal phonons the vacuum energy is

---

* Compare with the discussion of phonons in Ch. 3; the momentum $\pi(x) = -\tilde p(x)$, where $\tilde p(x) = \rho(x) - \rho_0$ is the density deviation from the mean ρ_0.

The conjugate field momentum to $A_\mu(x)$ for the EM field is of course just $\pi_\mu(x) = \delta \mathcal{L} / \delta \dot A^\mu(x)$. In the Coulomb gauge $\pi_0(x) = 0$, and $\underline{\pi}(x) = \underline{E}(x)$, the electric field; the Hamiltonian is $\mathcal{H} = \tfrac{1}{2} \int d^3r\, [E^2(r) + B^2(r)]$.

precisely that in eqtn (3), with $\lambda = 0$ only. It is a simple exercise to evaluate E_0^ϕ in this case; we do it at finite temperature to separate out E_T^ϕ from the total E. For definiteness we assume the Debye model, for which one simply imposes a sharp cut-off in the phonon density of states $N_\phi(\omega)$ at an energy $\hbar\Omega_D$, where Ω_D is the Debye frequency; writing

$$\sum_q \equiv \int d\omega\, N_\phi(\omega) \tag{21}$$

we have, in 3d, that

$$\left.\begin{aligned} N_\phi(\omega) &= 9\frac{n_0}{\Omega_D^3}\omega^2\,\theta(\hbar\Omega_D - \omega) \\ &\equiv 9 n_0 \left(\frac{\hbar}{k\Theta_D}\right)^3 \omega^2\,\theta(k\Theta_D - \hbar\omega) \end{aligned}\right\} \tag{22}$$

where $k_B\Theta_D \equiv \hbar\Omega_D$, where Θ_D is the Debye temperature, and n_0 is the number density of the system.

In thermal equilibrium, we have from (20) that

$$\left.\begin{aligned} E &= E_0^\phi + \sum_q \hbar\omega_q \frac{1}{e^{\beta\hbar\omega_q} - 1} \\ &\equiv E_0^\phi + 9 n_0 kT \left(\frac{T}{\Theta_D}\right)^3 \int_0^{\Theta_D/T} dx\, \frac{x^3}{e^x - 1} \end{aligned}\right\} \tag{23}$$

where the vacuum energy is

$$\left.\begin{aligned} E_0^\phi &= \frac{9}{8} n_0 k_B \Theta_D \\ &\equiv \frac{9}{8} n_0 \hbar\Omega_D \end{aligned}\right\} \tag{24}$$

Now, crucially, we CANNOT subtract this energy off. It plays no role in the thermal properties of the liquid - but the moment we try to impose any changes on the $T = 0$ system, we realise that E_0^ϕ plays a role. Consider, eg., the compressibility of the zero-T system, given by

$$\kappa = -\frac{1}{V}\left.\frac{\partial V}{\partial P}\right|_N \equiv \frac{1}{V}\left(\frac{\partial^2 E_0^\phi}{\partial V^2}\right)\Bigg|_N \tag{25}$$

since the pressure $P = -\partial E_0^\phi / \partial V\,|_N$; here we hold the total number N of particles constant. It is obvious that E_0^ϕ will change under a change of volume, ie., a change in density, since in (24), both n_0 and Ω_D will increase as P is increased; the result is readily calculated from (24) and (25). We see that there is a PRESSURE in the system coming from the vacuum zero-point energy, and it has a direct effect on properties like the

compressibility. It then follows that we cannot drop this term.*

Since we have argued that we can drop it for photons, the obvious question is - why the difference? Why is it, indeed, that it is considered quite permissible, in the case of a relativistic field like the EM field, with no UV cut-off, to use a <u>normal ordering procedure</u>, in which a product of operators like $\hat{O}_1\hat{O}_2$, which is some set of creation and annihilation operators, is re-ordered so as to put all creation operators on the left, and all annihilation operators on the right, ignoring all commutators. Thus, for our bosonic Hamiltonian $\mathcal{H}_0^{EM}$ in (16), we have

$$\mathcal{H}^0_{EM} = \frac{\hbar}{2}\sum_\lambda \int d^3q\, \omega_q \left[a^\dagger_\lambda(q)\, a_\lambda(q) + a_\lambda(q)\, a^\dagger_\lambda(q)\right] \tag{26}$$

but

$$\left.\begin{aligned} :\mathcal{H}^0_{EM}: &= \frac{\hbar}{2}\sum_\lambda \int d^3q\, \omega_q :\left[a^\dagger_\lambda(q)\, a_\lambda(q) + a_\lambda(q)\, a^\dagger_\lambda(q)\right]: \\ &= \hbar \sum_\lambda \int d^3q\, \omega_q\, a^\dagger_\lambda(q)\, a_\lambda(q) \end{aligned}\right\} \tag{27}$$

for the normal ordered form; the same technique is to produce any normal-ordered product $:\hat{O}_1\hat{O}_2:$ from the original $\hat{O}_1\hat{O}_2$. We see that the normal-ordering gets rid of the zero point energy, but it would certainly not be physically meaningful, according to our above argument, for the phonon system.

It turns out that this a subtle question, and we shall be returning to it several times again, notably:

- we will discuss how even if we try to drop the vacuum energy, we will still have to consider modifications of it when we distort the system, or change the boundary conditions (eg., by changing the volume by compression, as above, or by introducing or changing boundaries, as in the Casimir effect discussed below).
- we will calculate and analyze the stress-energy tensor, which is supposed to measure the energy, amongst other things; this will also be done when we have <u>anomalies</u> appearing in $T_{\mu\nu}(x)$, which can seriously complicate things.
- We will examine all the renormalization techniques used to deal with infinities in both $T_{\mu\nu}$ and all other physical quantities (mainly in Ch. 7).

One thing we will not do, unfortunately, is examine this question from the point of view of gravity. Here one runs into a very serious problem - the stress-energy tensor now acts as a source for spacetime curvature which in turn self-gravitates, ie., acts as a source for itself. Thus one has to deal now with infinities in both the gravitational vacuum energy (ie., in

vacuum energy associated with gravitons - this is quite apart from any singular configurations of the background classical field) but also the divergent vacuum of any field with its associated $T_{\mu\nu}$; and the interplay between the two. What works naively in flat space no longer works here — indeed, even if one subtracts off divergences of $T_{\mu\nu}$ in one configuration of spacetime (eg., in Minkowski space), it is found that the shift in the vacuum energy produced by changing this spacetime is in general infinite.

Thus we have a much more fundamental problem, which afflicts QFT even if we simply treat the background metric classically, ignoring any back reaction of the matter field fluctuations on it. Various recipes have been devised to deal with these problems, but they only work under restricted conditions, and quantum gravity coupled to matter is fundamentally non-renormalizable.

These questions are discussed in another set of course notes, on quantum gravity

DIRAC VACUUM : Similar ambiguities arise when we look at the Dirac vacuum; and just as with the EM field, it is helpful to look at a condensed matter analogue.

In what follows it will be helpful to look at the Appendix on Dirac fermions for details of the manipulations. Recalling the key points from these, we assume a Bogolubov decomposition of the fermion operators into electron & positron fields, viz.,

$$\left.\begin{aligned} \psi(x) &= \sum_{\underline{k}} \left(\frac{m}{E_k}\right)^{1/2} \sum_{\sigma} \left[b_{k\sigma}\, u_{k\sigma}\, e^{-ikx} + d^{\dagger}_{k\sigma}\, v_{k\sigma}\, e^{ikx} \right] \\ \bar{\psi}(x) &= \sum_{\underline{k}} \left(\frac{m}{E_k}\right)^{1/2} \sum_{\sigma} \left[b^{\dagger}_{k\sigma}\, \bar{u}_{k\sigma}\, e^{ikx} + d_{k\sigma}\, \bar{v}_{k\sigma}\, e^{-ikx} \right] \end{aligned}\right\} \quad (28)$$

where $b^{\dagger}_{k\sigma}$ creates an electron with spin σ, and $d^{\dagger}_{k\sigma}$ creates a positron with spin σ; thus we have the anticommutation relation

$$\{ b_{k\sigma}, b_{k'\sigma'} \} = \delta_{\underline{k}-\underline{k}'}\, \delta_{\sigma\sigma'} \quad (29)$$

and so on: and as usual, $\sum_{\underline{k}} \equiv \int d^3k/(2\pi\hbar)^3$. The Hamiltonian is then, as usual

$$\mathcal{H} = \pi_\alpha \partial_0 \psi_\alpha - \mathcal{L} = \bar{\psi}(x)\left[i\gamma^\mu \partial_\mu + m \right]\psi(x) \quad (30)$$

so that the total energy, integrated over space, is

$$\left.\begin{aligned} H = \int d^3x\, \mathcal{H}(x) &= i \int d^3x\, \bar{\psi}(x)\, \gamma^0 \partial_0\, \psi(x) \\ &= \sum_{\underline{k}} \sum_{\sigma} E_k \left[b^{\dagger}_{k\sigma} b_{k\sigma} - d_{k\sigma}\, d^{\dagger}_{k\sigma} \right] \end{aligned}\right\} \quad (31)$$

Now we are faced with a choice. We can either compute this directly,

by switching the operators in the 2nd term, so that we have

$$\left.\begin{aligned} H &= \sum_k \sum_\sigma E_k \left[b^\dagger_{k\sigma} b_{k\sigma} + \left(d_{k\sigma} d^\dagger_{k\sigma} - 1 \right) \right] \\ &\equiv \sum_k \sum_\sigma E_k \left[b^\dagger_{k\sigma} b_{k\sigma} + d^\dagger_{k\sigma} d_{k\sigma} \right] + E_0^\psi \end{aligned}\right\} \quad (32)$$

where the vacuum energy is: $$E_0^\psi = -\sum_{k,\sigma} E_k \equiv -\int \frac{d^3k}{(2\pi)^3} E_k \quad (33)$$

per unit volume.

Thus we find a vacuum energy which is both negative and infinite; this is apparently even worse than what we find for a bosonic field! On the other hand we can normal order the expression in (31); for fermions this means we write

$$\left.\begin{aligned} :d_{k\sigma} d^\dagger_{k\sigma}: &= :\left[1 - d^\dagger_{k\sigma} d_{k\sigma} \right]: \\ &= - d^\dagger_{k\sigma} d_{k\sigma} \end{aligned}\right\} \quad (34)$$

at which point the vacuum energy becomes zero, i.e., in sloppy notation we have

$$:E_0^\psi: \;=\; 0. \quad (35)$$

Now which of these two is correct? In the context of relativistic QFT, this depends on what theoretical framework you are using. In ordinary QFT in flat spacetime, most authors choose (35), so that $E_0^\psi = 0$. On the other hand in a supersymmetric theory, the negative fermionic vacuum energy cancels with its bosonic superpartner, to give net zero, without any normal ordering or UV cutoff. For many this is a powerful argument for supersymmetry.

Further light on this is cast if we go to a condensed matter analogue. Since the dimensionality of the theory is not a key part of the above derivation, the simplest thing is to choose a 1+1-dimensional system; a very nice example (not the only one of course) is the polyacetylene system, for which the Hamiltonian can be written as

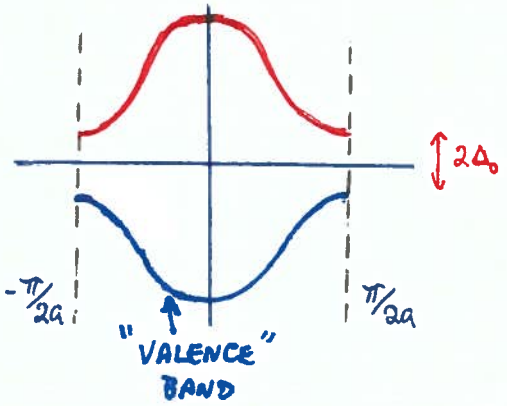
$$\left.\begin{aligned} H = \int_{-\infty}^{\infty} dx \sum_\sigma \psi^\dagger_\sigma(x) \left[-i\hbar v_F \hat{\sigma}_3 \partial_x + \Delta(x) \hat{\sigma}_1 \right] \psi(x) \\ + \frac{\Omega_0^2}{g^2} \int_{-\infty}^{\infty} dx \, \Delta^2(x) \end{aligned}\right\} \quad (36)$$

the origin of which is discussed more thoroughly in an Appendix. Here $\psi(x)$ describes electrons, and $\Delta(x)$ is a "lattice displacement field", in the continuum limit, which can not only fluctuate (essentially as an optical phonon field), but also acquire "kink" soliton and anti-soliton configurations. The energy Ω_0^2 is an optical phonon energy, and g^2 is a dimensionless electron-phonon coupling.

For our purposes, we assume that $\Delta(x) \to \Delta_0$, a constant, which then plays the role of a mass in the Dirac eqtn; we thus ignore the dynamical phonon field, and the soliton excitations.

We can then go through the same calculation of the vacuum energy as we did for QED, and we get essentially the same result as in in (33); one finds

$$E_0^{\psi} = \int_{-\pi/2a}^{\pi/2a} \frac{dk}{2\pi\hbar} \sum_{\sigma} E_k f_k = \lim_{\beta\to\infty} \int_{-\pi/2a}^{\pi/2a} \frac{dk}{2\pi\hbar} \sum_{\sigma} E_k \frac{1}{e^{\beta(E_k-\mu)}+1} \qquad (36)$$



where the band structure is as shown in the figure. Now the answer we get for (36) depends entirely on where we put the chemical potential. If, as will be the case if the system is undoped, we put μ in the middle of the band gap between the lower "valence" band and the upper band, then we can rewrite (36) as a sum over only the negative energy valence band states, and we find a negative vacuum energy; this is obvious from the figure, since all occupied valence band states have negative energy. Formally we have

$$\left.\begin{aligned} E_k^2 &= \epsilon_k^2 + \Delta_0^2 \\ \epsilon_k &= 2t_0 \cos ka \end{aligned}\right\} \qquad (37)$$

where a is the lattice spacing, and t_0 is an intersite hopping energy; we then just have

$$E_0^{\psi} = -\sum_{\sigma} \int_{\Delta_0}^{2t_0} \frac{dE}{2\pi\hbar} N(E) E \qquad (38)$$

where $N(E_k) = dk/dE_k$ is the density of states.

In this case it is obvious what is going on - the filled states all have negative energy with respect to μ, so the vacuum must have negative energy. We see that in QED, the naive evaluation of E_0^{ψ}, to give (33), corresponds precisely to this. On the other hand if we normal order the operators, it is as though we have shifted μ down in energy to cancel the result exactly. Recall our discussion of Fermi liquids (cf. eqtns (7) and (9)).

From all this discussion it will be clear that some care, as well as physical insight, is required to properly understand this issue.

INTERACTIONS in QED : We only make a few remarks here, because the technical details require us to go into some depth. The following points are key:

- Although the coupling constant is small, nevertheless a series expansion for the effect of the electron-photon coupling on the vacuum energy is formally divergent, even after regularization / renormalization. This is an almost ubiquitous feature of perturbation expansions in QFT. However, if we view the series as an asymptotic series, then it makes mathematical sense, and gives very accurate answers for physical quantities.

  We discuss the asymptotic properties of both scalar field theory and of QED below, in discussion of vacuum decay - this will allow us to look at large orders in perturbation theory.

- One can study a theory as a function of complex coupling coefficient - this turns out to be a very useful method for looking at the asymptotic expansion noted above. One then finds interesting structure in the complex coupling constant plane, including branch cuts (corresponding to vacuum instabilities), for the vacuum energy.

- Finally, we can ask what happens to the vacuum in QED when we change either the boundary conditions on the vacuum (eg., by imposing boundaries, or more exotically, with a background spacetime curvature), or if we couple to some static external current (eg., by applying a strong electric or magnetic field, coming from a charge or electric current). One can of course try to answer these questions perturbatively, but as will see, the non-perturbative effects turn out to be extremely important and interesting.

(iii) PATH INTEGRAL FORMULATION :

Up to now we have done only simple calculations designed to elucidate the properties of the ground state energy & the vacuum; and we have alluded to the effects of interactions by looking at perturbation theory and low-order diagrams. However neither of these is terribly helpful in elucidating the true non-perturbative properties of the vacuum.

To go further we must go back to our functional formulation of QFT; and since we are interested in the ground state, we must go back to the vacuum generating functional $Z[J]$, with J set to zero. In keeping with the general tenor of this section, we will emphasize the properties of simple models - which turns out to be good, since their properties are already rather

complicated. The discussion will be divided between fermionic & bosonic theories, since the path integral takes a different form for each of these.

- **INTERACTING SCALAR FIELD** : For our bosonic example we choose the standard scalar field theory, ie., the generating functional of $J = 0$ will be

$$Z[J=0] = N_0 \int \mathcal{D}\phi \, e^{i/\hbar \int d^D x \left\{ \frac{1}{2}[\partial_\mu \phi \partial^\mu \phi - m^2 \phi^2] - g/4! \phi^4 \right\}} \tag{39}$$

where N is a normalization factor, which is of course infinite. In previous discussions, where $J(x)$ was non-zero, we simply set N to be equal to $Z[J=0]$, so that the vacuum transition amplitude was unity. Here, however, we can't do this, because we wish to study the properties of $Z[J=0]$, so we need to find another way to regularize the path integral.

One simple way, which is related to Pauli-Villars regularization, is to simply define N in terms of a new field with a different mass M; thus we will write

$$N^{-1} = \int \mathcal{D}\phi \, e^{i/\hbar \int d^D x \left\{ \frac{1}{2}[\partial_\mu \phi \partial^\mu \phi - M^2 \phi^2] - g/4! \phi^4 \right\}} \tag{40}$$

when necessary; we will then let $M \to 0$ at the end.

We now wish to determine the vacuum energy for the theory in eqtn (39). To do this we recall that the vacuum amplitude can also be related to the vacuum energy of a field theory; one has *

$$Z[J=0] = \lim_{\substack{T \to \infty \\ T' \to -\infty}} \langle 0 | e^{-i/\hbar \, \mathcal{H}(T-T')} | 0 \rangle \tag{41}$$

and by taking the log of this, we have

$$\ln Z[J=0] = -i/\hbar \, E_0^\phi (T-T') \tag{42}$$

since only the vacuum contribution to $\mathcal{H}$ is then picked out.

(a) **NON-INTERACTING CASE** : We start by assuming that $g = 0$ in the action in eqtn. (39); this means we will be retaining the same free field vacuum

---

* For a thorough discussion of this, see the notes on "Correlation functions in Quantum Mechanics" in section A of these lectures.

as we already found (ie, the sum of zero-point energies for an infinite set of oscillators). From (39) and (40) we have

$$Z_0[J=0] = \frac{\int \mathcal{D}\phi\, e^{i/\hbar \int d^D x\ \frac{1}{2}[\partial_\mu\phi\partial^\mu\phi - m^2\phi^2]}}{\int \mathcal{D}\phi\, e^{i/\hbar \int d^D x\ \frac{1}{2}[\partial_\mu\phi\partial^\mu\phi - M^2\phi^2]}}\Bigg|_{M=0} \tag{43}$$

where we will set $M=0$ at the end. From (43) we immediately find that

$$Z_0[J=0] = \left(\frac{\det|\partial^2+M^2|}{\det|\partial^2+m^2|}\right)^{1/2\hbar}_{M=0} \equiv \frac{e^{-\frac{1}{2}\,\mathrm{Tr}\,\ln(\partial^2+m^2)}}{e^{-\frac{1}{2}\,\mathrm{Tr}\,\ln(\partial^2+M^2)}}\Bigg|_{M=0} \tag{44}$$

where the Tr simply means $\int d^D x\ \ln(\partial^2+m^2)$. Fourier transforming, and taking the log of both sides, we get

$$\left.\begin{aligned} i\hbar\, Z_0[J=0] &= -\frac{i\hbar}{2}\int d^D x \sum_q \ln\left|\frac{q^2-m^2-i\delta}{q^2-M^2-i\delta}\right|_{M=0} \\ &= E_0(T-T') \end{aligned}\right\} \tag{45}$$

where in the last step we used (42); and $\sum \equiv \int d^D q/(2\pi\hbar)^D$. Now the spacetime integral $\int d^D x$ is simply equal to the product of the spatial volume V and the infinite time interval $T-T'$, since the integrand is independent of x; thus we also have

$$\frac{E_0}{V} = -\frac{i\hbar}{2}\sum_q \ln\left|\frac{q^2-m^2-i\delta}{q^2-M^2-i\delta}\right|_{M=0} \tag{46}$$

for the energy per unit volume of the vacuum. Separating out the frequency integral, this just becomes, per unit volume, the vacuum energy:

$$\left.\begin{aligned} \mathcal{E}_0^{\phi} &= -\frac{i\hbar}{2}\sum_{\underline{k}}\int_{-\infty}^{\infty}\frac{dk_0}{2\pi\hbar}\ln\left|\frac{k_0^2-(\underline{k}^2+m^2)-i\delta}{k_0^2-(\underline{k}^2+M^2)-i\delta}\right|_{M\to 0} \\ &= i\hbar\sum_{\underline{k}}\int_{-\infty}^{\infty}\frac{x\,dx}{2\pi\hbar}\left[\frac{x}{x^2-\omega_k^2-i\delta}-\frac{x}{x^2-\Omega_k^2-i\delta}\right]\Bigg|_{\Omega_k\to 0} \end{aligned}\right\} \tag{47}$$

where we set $\omega_k^2 = \sqrt{|\underline{k}|^2+m^2}$, $\Omega_k^2 = |\underline{k}|^2+M^2$, and integrate by parts. This integral is simple, and it gives

$$E_0^\phi = \lim_{\Omega_k \to 0} \frac{\hbar}{2} \sum_{\underline{k}} \omega^0_{\underline{k}} - \Omega_{\underline{k}} = \frac{\hbar}{2} \sum_{\underline{k}} \omega^0_{\underline{k}} \qquad (48)$$

as we expect. We see that the regulator mass M simply serves to make the integrals well-defined, and can be dropped at the end, as also expected.

(b) INTERACTING CASE : We now turn to the much more elaborate problem of an interacting scalar field. There are a number of key things that need to be discussed here, notably

- The relationship between perturbation theory and the non-perturbative results that one can derive by other means
- The possibility that the vacuum may be unstable - this cannot be discussed using perturbation theory except in an indirect way, and it leads one to the discussion of both spontaneously broken symmetries, and to the "instanton" formalism associated with vacuum decay.
- Even if the vacuum is not unstable, one may still have to take account of tunneling processes between different degenerate vacua of the classical theory.

Here we will only discuss the 1st topic in any detail - we save the discussion of tunneling, vacuum decay, and instantons until we get to this topic. What is meant here is easily seen in the discussion of the ϕ^4/ϕ^6 model (compare the action in (1)), with the effective potential curves shown in the figure after eqtn (1)). If $m^2, g_4, g_6 > 0$, we may hope to extract results using perturbation theory in a simple way; but when $m^2 < 0$, we get degenerate minima, and if $g_4 < 0$, we get a metastable vacuum at $\phi(x) = 0$; the actual vacuum has finite $|\phi|$. If $g_6 < 0$, the theory is completely unstable.

Thus we will look here only at the 1st case, where $m^2, g_2, g_4 > 0$. The question then is — what can perturbation theory tell us? Consider again the generating functional with $J = 0$, viz.,

$$Z[J=0] = \mathcal{N} \int D\phi \; e^{\frac{1}{\hbar} \int d^D x \left\{ \frac{1}{2}(\partial_\mu \phi \partial^\mu \phi - m^2 \phi^2) - \frac{g_4}{4!}\phi^4(x) - \frac{g_6}{6!}\phi^6(x) \right\}} \qquad (49)$$

where we will not need to deal with the normalizing factor $\mathcal{N}$. Again, we want to calculate the ground state energy $E_0^\phi(g_4, g_6)$, this time as a function of g_4 and g_6; we want to know $\Delta E_0 = E_0^\phi(g_4, g_6) - E_0^\phi(0,0)$, where we just calculated $E_0^\phi(0,0)$ in (48) above.

Just to vary things a little bit, instead of using (45) to find E_0, we rotate to imaginary time; then we can define the "free energy"

$$F = -1/\beta \ln [Tr\, e^{-\beta H}] \tag{50}$$

where $\beta = 1/kT$ play the role of imaginary time (so we will take $\beta \to \infty$), and $\mathcal{H}$ is the "imaginary time action", given here by

$$\mathcal{H} = -\int d^{D-1}r \int_0^\beta d\tau \left\{ \tfrac{1}{2}(\partial_\mu \phi \partial^\mu \phi + m^2\phi^2) + \frac{g_4}{4!}\phi^4(r,\tau) + \frac{g_6}{6!}\phi^6(r,\tau) \right\} \tag{51}$$

so that the energy shift is.

$$\Delta E^\phi = E_0^\phi(g_4, g_6) - E_0^\phi(0,0) = \lim_{\beta\to\infty} -1/\beta \ln\left[\frac{Z[g_4, g_6; J=0]}{Z_0[0,0; J=0]}\right] \tag{52}$$

Now the key contribution of a group of people* working in the 1970's, building on an earlier fundamental idea due to Langer**, was to recognize that the evaluation of these functional integrals was a fairly straightforward generalization of the problem for either a 1-d or a zero-d anharmonic oscillator, and that it could be done using standard steepest descent methods. To see why this may be, consider the following simple integral:

$$I(g) = \int_{-\infty}^{\infty} dx\, e^{-(\frac{1}{2}x^2 + g/4 x^4)} \tag{53}$$

We can evaluate this as a power series expansion, or using a saddle point integration; it can also be evaluated exactly, in terms of Whittaker functions (or equivalently, in terms of parabolic cylinder functions, or Bessel/Hankel functions, since these are all related). As a power series, we write

$$I(g) = \sum_{n=0}^{\infty} C_n g^n \tag{54}$$

where $$C_n = \frac{1}{n!}\left(\frac{-1}{4}\right)^n \int_{-\infty}^{\infty} dx\, e^{-x^2/2} x^{4n} = \frac{(2\pi)^{1/2}}{16^n}(-1)^n \frac{(4n)!}{n!\,(2n)!} \tag{55}$$

---

* For work on the functional treatment of ϕ^{2n} theory, and the connection to high-order perturbation theory, see: C.M. Bender, T.T. Wu; in PRL 37, 117 (1976), following on from Phys. Rev. 184, 1231 (1969), P.R. D7, 1620 (1973); T. Banks, C.M. Bender, TT Wu, Phys. Rev. D8, 3346 (1973), T. Banks, CM Bender. Phys. Rev. D8, 3366 (1973); L.N. Lipatov, JETP Lett. 24, 157 (1976), and 25, 104 (1977); and JETP 44, 1055 (1976); 45, 216 (1977); E. Brezin, JC le Guillou, J Zinn-Justin, Phys. Rev. D15, 1544 & 1588 (1977); etc.

** J.S. Langer, Ann. Phys 41, 108 (1967); see also C.S. Lam, Nuovo Cimento 55A, 258 (1968)

and for large n, using Stirling's approximation, we have

$$c_n \xrightarrow[n>>1]{} \left(\frac{2}{n}\right)^{1/2}\left(\frac{-4n}{e}\right)^n \qquad (56)$$

Thus we deal with an asymptotic series which, as $n \to \infty$, blows up (with alternating sign for even/odd n), and which approaches $I(g)$ most closely when $n \sim O(1/g)$, for $g<<1$.

The same result is obtained by a steepest descent treatment of the integral in (53); we have

$$\left.\begin{aligned} c_n = \frac{1}{n!}\left(\frac{-1}{4}\right)^n \int_{-\infty}^{\infty} dx\, e^{-x^2/2}\, e^{4\ln n} &\underset{n>>1}{\sim} \frac{2\sqrt{\pi}}{n!}\left(\frac{-1}{4}\right)^n n^{2n} e^{-2n(1-\ln 4)} \\ &\sim (-1)^n \frac{\sqrt{2}}{n}\, e^{n\ln 4n - n} \end{aligned}\right\} \qquad (57)$$

where we have a saddle point at $x \to x_0$, with $x_0^2 = 4n$, and the integral is evaluated up to quadratic terms around the saddle point. Eqtns (56) and (57) are consistent with each other. For those who are interested, the exact result for the integral is

$$\left.\begin{aligned} I(g) &= 2^{-3/2}\, g^{-1/2}\, e^{1/8g}\, K_{1/4}(1/8g) \\ &\equiv i\pi 2^{1/2}\, e^{i\pi/8}\, g^{-1/2}\, e^{1/8g}\, H^{(1)}_{1/4}(e^{i\pi/2}/8g) \\ &\equiv \frac{2^{-3/2}}{\pi}\, g^{1/2}\, e^{1/8g}\, W_{0,1/4}(1/4g) \end{aligned}\right\} \qquad (58)$$

where $K_\nu(z)$, $H^{(1)}_\nu(z)$, and $W_{0,\nu}(z)$ are imaginary argument Bessel functions, Hankel functions, and Whittaker functions respectively. Note that the result in (58) is actually only valid under the condition, if we view g as a complex variable, that $\mathrm{Re}\, g > 0$. There is a good reason for this — if $\mathrm{Re}\, g$ is negative, the integral blows up.

One cannot immediately see this problem by looking at the asymptotic series, because the successive terms also diverge as $n \to \infty$, no matter what g is doing. It does, however, become obvious if we sum the divergent series. This can be done by Borel transformation of the integral. The details of this are given in the appendix; here one sums the series

$$B(g) = \sum_{n=0}^{\infty} c_n/n!\; g^n \qquad (59)$$

instead of $I(g)$ in (54); if this converges, the inverse Borel transform

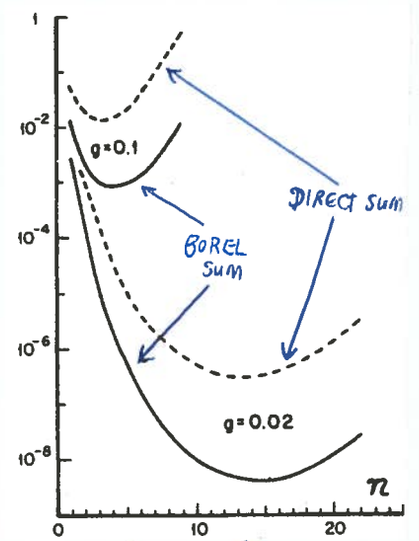
is the

$$I_B(g) = \int_0^\infty dx\, e^{-x} B(gx) \tag{60}$$

and if the original series is not too divergent, $I_B(g) \to I(g)$. This can be done with the series in (56) or (57), and the result is

$$I_B(g) = -\int_0^\infty \frac{dx}{\pi^{1/2}}\, e^{-x} \ln(1+4gx) \tag{61}$$

which has a branch cut along the real axis when $\mathrm{Re}\, g < -1/4$. As we will discuss later, we can analyze all of this behaviour using an instanton approach, which directly treats the tunneling process involved for negative g.



PLOTS OF $\left|I(g) - \sum_{r=1}^{n} C_r g^r\right|$, FOR 2 DIFFERENT VALUES OF g, EVALUATED BY DIRECT SUM AND BOREL SUMMATION.

What do we learn from all of this? The first thing is we see very clearly the asymptotic nature of any Feynman graph summation, even for a perfectly sensible theory - as terms go to high order, the series diverges, and the best estimate for the actual answer appears at finite order, for $n \sim O(1/g)$, where g is the relevant coupling. This is shown at left for $I(g)$. If $g \ll 1$, then we expect to get very accurate answers using this method.

Second, we see that we can learn a lot about $E_0(g)$ by analyzing the behaviour of the high-order perturbative contributions to $Z[g]$; this is because all the considerations we have just discussed for the "$d=0$" field theory here (ie, the integral $I(g)$), go over to the full scale field theory, if we employ a saddle-point / steepest descent approximation to evaluate the functional integral in (49);

ie., we write

$$N \int \mathcal{D}\phi\, e^{\frac{i}{\hbar} S[\phi]} = N e^{\frac{i}{\hbar} S[\phi_c(x)]} \int \mathcal{D}\phi\, e^{\frac{i}{2\hbar} \iint \frac{\delta^2 S}{\delta\phi(x)\,\delta\phi(x')}\Big|_{\phi_c} \delta\phi(x)\,\delta\phi(x')} \tag{62}$$

where $\delta\phi(x) = \phi(x) - \phi_c(x)$, and $\phi_c(x)$ is the (semiclassical) solution for $\phi(x)$ when $S[\phi]$ is at a saddle point.

To go through the details of this calculation is rather tedious - the reader is referred to the papers of Lipatov, and of Brezin et al., cited above.

- **FERMIONIC THEORIES** : I will say a lot less here, since the basic ideas are the same as for the scalar field discussed above, but the execution of these ideas is usually much messier and more technical. Let's

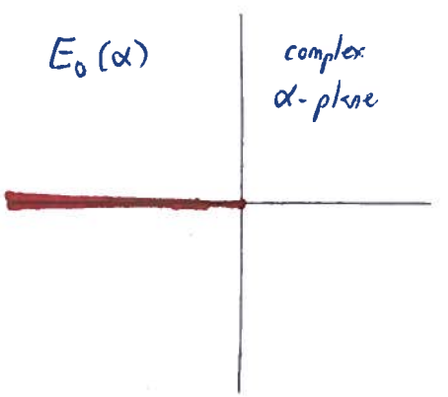
look first at QED, which has been well-studied, and then say a little about other systems.

For QED, one key physical point was made clearly by Dyson* in 1952; it is simple and yet of great generality. Consider a quantity like the vacuum energy in QED, written as a power series expansion in the dimensionless coupling g^2 (where in QED, $g^2 = \alpha = e^2/$ 1/137 in "natural units" where $\hbar = c = \varepsilon_0 = 1$, as usual**). We write

$$E_0(\alpha) = E_0(\alpha=0) + \sum_{n=1}^{\infty} C_n \alpha^n \qquad (63).$$

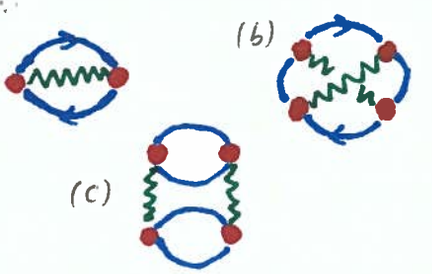
Note that here we are talking about the vacuum energy - but we could equally well be expanding any other physical quantity in a similar power series.

Now if (63) is meaningful, $E_0(\alpha)$ must be an analytic function of α, at least for $|\alpha| < \alpha_0$, where α_0 is the radius of convergence about the origin. However a simple physical argument suffices to show that $\alpha_0 = 0$. Suppose we let $\alpha \to -\alpha$, i.e., in the complex α-plane, we go onto the negative real axis. Physically this corresponds to making the electron-positron interaction repulsive, and the electron-electron (or positron-positron) interaction attractive.



ANALYTIC STRUCTURE OF $E_0(\alpha)$ IN QED, SHOWING THE BRANCH CUT ALONG THE NEGATIVE α-AXIS.

What would then happen to the vacuum under these circumstances? Consider the simplest processes, in low-order perturbation theory, that contribute to $E_0(\alpha)$; these include graphs like those shown.



Now if $\alpha > 0$, we expect each of these diagrams to raise the energy of the vacuum - it takes energy to separate the electron-positron pairs. However when $\alpha < 0$, the opposite is true; the system can LOWER its energy by spontaneously creating electron-positron pairs.

In practice this is a tunneling phenomenon: there is a mass gap of $2m_e c^2$ to overcome, and yet the repulsive electron positron interaction goes like $V(r) = \pi\alpha/r$. Thus there is a potential barrier blocking the escape, and

---

* F.J. Dyson, Phys. Rev. 85, 631 (1952)

** In SI units, $\alpha = \frac{1}{4\pi\varepsilon_0} e^2/\hbar c \equiv \frac{\mu_0}{4\pi} e^2 c/\hbar$. In cgs units, $\alpha = e^2/\hbar c$. If we let $\varepsilon_0 = \hbar = c = 1$ in the SI units, then $\alpha = e^2/4\pi$.

we infer that there is a branch cut in the complex α-plane, of magnitude

$$\mathrm{Im}\, E_0(\alpha) \sim O\left(e^{-C_0/|\alpha|}\right) \qquad (\mathrm{Re}\,\alpha < 0,\ \mathrm{Im}\,\alpha = 0). \tag{64}$$

where $C_0 \sim O(1)$. To properly study this involves instanton methods, which we defer to later on; these are basically an adaption of the high-order perturbation methods referred to before, to the QED case, in an unstable regime similar to that described above for scalar field theory.*

Let's briefly note how one deals with this formally, using a path integral formulation. We start from the sourced QED generating functional, which as usual takes the form

$$Z_{QED}[J^\mu, \bar\eta, \eta] = N \int \mathcal{D}A_\mu\, \mathcal{D}\bar\psi \mathcal{D}\psi\, e^{\frac{i}{\hbar}\left(S_{QED} + \int d^4x\,[J^\mu A_\mu + \bar\psi\eta + \bar\eta\psi]\right)} \tag{65}$$

and now let's integrate out the fermionic degrees of freedom, and then set the external fermionic currents to zero. We did this same exercise for coupled scalar fields in Chapter 3; the technique is the same, except we now deal with fermionic determinants. The result is

$$\left.\begin{aligned} Z_{QED}[J^\mu, \bar\eta, \eta] = N \int \mathcal{D}A_\mu\, & e^{\frac{i}{\hbar}\left(S_0[A_\mu] + \int J^\mu A_\mu\right)} e^{-\frac{i}{\hbar}\iint \bar\eta\, \tilde{\mathcal{G}}[A_\mu]\, \eta} \\ & \times e^{-\mathrm{Tr}\,\ln\left(\tilde{\mathcal{G}}_2[A_\mu]/G_0\right)} \end{aligned}\right\} \tag{66}$$

where the propagator $\mathcal{G}[A_\mu] \equiv \mathcal{G}(x,x'|A_\mu)$ is that for the Dirac fermion moving in a "frozen" $A_\mu(x)$ field, which we already met in the last section, satisfying

$$\left[\gamma^\mu(i\partial_\mu + eA_\mu) - m\right] G_2(x,x'|A_\mu) = -\delta(x-x') \tag{67}$$

and where $S_0[A_\mu]$ is just the free photon action. Notice the key difference in the last term in (66) from what we found for coupled scalar fields in Ch. 3 (compare eqtn. (113) and (126) in Ch. 3); here we have

$$\exp\left\{-\mathrm{Tr}\,\ln\left(\frac{\tilde{\mathcal{G}}_2(x,x'|A_\mu)}{G_0(x,x')}\right)\right\} \equiv \exp \mathrm{Tr}\,\ln\left[1 + e\gamma^\mu A_\mu(x)\, G_0(x,x')\right] \tag{68}$$

---

* Here we have ignored other possible instabilities that can arise in QED when the interactions are strong – one can then in principle get a phase transition to a "confined" phase, in which the electrons & positrons bind – the vacuum is then quite different.

which we can compare with Ch. 3, eqtn (114) (and noting that there is no square root in the determinant now, because we have integrated over both $\bar{\psi}$ and ψ); here, as usual, we have

$$\left.\begin{aligned} G_0(x,x') &= i\hbar\, S_F(x,x') = i\hbar\left(i\gamma^\mu\partial_\mu - im + i\delta\right)^{-1} \\ &= \sum_k e^{ik(x-x')}\,\frac{i\hbar}{\gamma^\mu k_\mu - m + i\delta} \end{aligned}\right\} \qquad (70)$$

If we now set the external currents J^μ, $\bar{\eta}$, and η to zero, we can write

$$\left.\begin{aligned} Z_{QED}(0,0,0) &= N\int DA_\mu\, e^{\frac{i}{4\pi}S_0[A_\mu]}\, e^{Tr\ln\left[1+e\gamma^\mu A_\mu(x) G_0(x,x')\right]} \\ &\equiv N\int DA_\mu\, e^{\frac{i}{4\pi}S_0[A_\mu]}\ \det\left|1+e\gamma^\mu A_\mu(x)\, G_0(x,x')\right| \end{aligned}\right\} \qquad (71)$$

where the normalizing factor N contains of course the generating functional for the non-interacting Dirac fermions. This is easily evaluated, in the same way as we did for the free scalar field above; we have

$$\left.\begin{aligned} Z_0^{\psi}[0,0,0] &= N_0 \int D\psi\, D\bar{\psi}\, e^{\frac{i}{\hbar}\int d^4x\, \bar{\psi}\left(i\gamma^\mu\partial_\mu - m + i\delta\right)\psi} \\ &= N_0\, e^{Tr\ln\left(i\gamma^\mu\partial_\mu - m + i\delta\right)} \end{aligned}\right\} \qquad (71)$$

where we can, if we wish, play the same trick in writing N_0 as we did for the scalar field before. We can even write the determinant / Tr log form in the same way as for the scalar field, by using the standard trick in which we multiply γ^μ to the left and right by γ^5; then $Tr\,\gamma^5\gamma^\mu\gamma^5 = Tr(\gamma^5)^2\gamma^\mu = Tr\,\gamma^\mu$. However, since $\gamma^5\gamma^\mu\gamma^5 = -\gamma^\mu$, we have $Tr\left(i\gamma^\mu\partial_\mu - m\right) = Tr\left(-i\gamma^\mu\partial_\mu - m\right)$, and so we can write $Tr\ln\left(i\gamma^\mu\partial_\mu - m\right) = \tfrac{1}{2}\ln\left[Tr\left(i\gamma^\mu\partial_\mu - m\right) + Tr\left(-i\gamma^\mu\partial_\mu - m\right)\right]$, so that

$$Z_0^{\psi}[0,0] = N_0\, e^{\frac{1}{2}Tr\ln\left(\partial^2 + m^2 - i\delta\right)} \qquad (72)$$

which has the same form as the free boson result in (44), except for the sign of the exponent. Following through in the same way as we did for eqtns (45)–(48), we just get

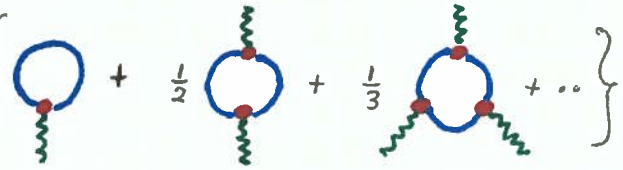
$$E_0^{\psi} = i\hbar\ln Z_0^{\psi}[0,0] = -\sum_{k,6} E_k \qquad (73)$$

where $E_k^2 = k^2 + m^2$, i.e., the same result as found earlier using a simple sum over

vacuum modes — see eqtn (33).

For the interesting case we see from (71) that we have the standard "loop sum" for the Tr log term; the only difference is a factor of (-1) coming from change of sign, as compared to the bosonic case. Thus we can represent the determinant in (71) as shown in the figure

$$\mathrm{Tr}\,\det\left|1+ey\!\!\!/A_\mu G_0(x,x')\right| \;=\; -\Big\{ \;+\; \tfrac{1}{2} \;+\; \tfrac{1}{3} \;+\;\cdots\Big\}$$



although it should immediately be noted that graphs with odd numbers of photons are zero (Furry's theorem). If we now exponentiate this determinant, we generate arbitrary numbers of loops, and then functionally integrating over A_μ joins all the photon lines in all possible ways — we then get all the usual diagrams for the vacuum amplitude. Taking the log, as in (73), then gives the vacuum energy.

The shift in the vacuum energy caused by the electron-photon interaction is of course infinite, because of UV divergences: we address this problem in Chapter 7. Once one has dealt with this issue, by properly regularizing the integrals over loops, we can in principle look at asymptotic expressions for the vacuum shift, in the same as discussed above for the bosonic field.

This problem has been studied since the late 1970's, with some interesting results.* Unfortunately it is considerably more difficult than the bosonic case; even the evaluation of the determinant is in general intractable,** and so one must resort to various approximation schemes. Things are a little easier if one reduces the number of dimensions (and in 1+1 dimensions QED goes over to the Schwinger model, to be studied in more detail in the next section — this is actually solvable). Thus, unfortunately, we do not know as much about the high-order terms, or about $E_0(\alpha)$, as we would like.

As we shall see, this fermionic determinant is of considerable importance — the first discussion of its properties goes all the way back to work of Heisenberg and Euler, and of Weisskopf, in 1936! If one freezes the configuration $A_\mu(x)$ in some simple form (eg., a static field, or a simple EM wave) one can go a lot further — indeed, perhaps the very first non-perturbative calculation in QED, by Schwinger in 1952, was of precisely a problem of this kind; we will examine this below, along with other related work.

---

* For early work, see C. Itzykson, G. Parisi, JB Zuber, PRL 38, 306 (1977); Phys Rev. D16, 996 (1977); R. Brout, C. Itzykson, JB Zuber, G. Parisi, Phys. Rev. D17, 1041 (1978); and SL Adler, Phys. Rev. D10, 2399 (1974). For non-Abelian gauge theories, see LN Lipatov, AP Bukhvostov, EI Malkov, Phys. Rev. D19, 2974 (1979)

** See, eg: GV. Dunne, J Phys. A41, 1 (2008)

B.6.3(b) THE PERTURBED QUANTUM VACUUM

In both condensed matter & relativistic quantum field systems, it is often more physically relevant to calc how the vacuum *changes* under perturbations - we then see very interesting behaviour which tells us a lot about the underlying vacuum. However, as we shall see, arguments persist about how one is to interpret these changes - either in terms of perturbed vacuum modes, or in terms of properties of the system which is perturbing them.

In what follows we will focus on 2 kinds of change. One is that brought about by an external field $J(x)$. You've by now aware of the importance of looking at the effects of a source current — now we will look directly at a key non-perturbative effect of this, often called the "Schwinger mechanism" (although it was first discussed 80 yrs ago by Heisenberg & Euler). This consists in looking at the slow "decay" of the vacuum brought about by a strong static or AC EM field. There are also interesting condensed matter analogues

We then discuss a topic of great interest in the last 30 yrs, viz., the "Casimir effect". This famous effect, first noted by Casimir in 1948, refers to what happens when we put some sort of "boundary" on the vacuum - eg, in a QED vacuum, by confining the vacuum by a set of conducting plates, or some dielectric medium. The importance of this effect is both practical - the world is full of material objects, which constrain the vacuum quite automatically - and theoretical, because it raises important questions about how to calculate these effects, and what they mean.

One thing I will not discuss here, although it falls into the same category, is the effect of "curved" backgrounds on quantum fields. Such questions vary from the quite prosaic (the effect of, eg, a background strain field on electron dynamics) to the fundamental (behaviour of an underlying curved spacetime on a quantum field). These issues are addressed (a) in the section on background field methods, and (b) in the course on quantum gravity.

(i) THE HEISENBERG-EULER-SCHWINGER EFFECT

We discuss this effect first in QED, and then briefly mention condensed matter analogues. The basic question here is — suppose we subject the QED vacuum to a constant external applied field, which could be some combination of $\underline{E}$ and $\underline{B}$ fields, what will the new vacuum look like?

Physically, we can see that this question is closely related to our discussion of the stability of the QED vacuum when the sign of α is reversed. If we apply a strong electric field, it will act oppositely on the electrons & positrons in a virtual electron-positron pair vacuum fluctuation. One can then imagine a "vacuum tunneling" process in which such a virtual pair is pulled apart

by the field. Clearly one can make an estimate of the probability per unit volume of spacetime for this to happen. Let's quickly do this, using a simple WKB argument. To create an electron-positron pair requires an energy $2m_e c^2$; the energy for this will come from the electric field pulling the pair apart. The pair is created in a region of typical size $\lambda = \hbar/mc$, the Compton wavelength, according to the uncertainty principle; the electric field energy involved is thus $eE\lambda = eE\hbar/mc$. This suggest a rate

$$\Gamma(x) = \Gamma_0 \sim e^{-2m^2c^3/e\hbar E_0} \qquad (74)$$

where the mass gap $2m_e c^2$ is acting like a barrier.

To give a better (indeed, an exact) non-perturbative result, we look at the classic work of Heisenberg & Euler, and of Schwinger, which solved this problem.

We begin by noting that the rate of pair production, in a <u>fixed</u> EM field $\bar{A}_\mu(x)$, is given very simply by looking at the |transition amplitude|^2, i.e., we can write, at a spacetime point x:

$$\frac{d}{dt}\left|\mathcal{G}(x,x\,|\bar{A}_\mu)\right|^2 = -\Gamma(x) \qquad (75)$$

Now note that:

(i) the vacuum amplitude $|Z_{QED}[J_\mu]|^2$ is given by the integral over spacetime of the probability $|G_{00}[J_\mu]|^2$, with the time limits taken to $\pm\infty$, and the inner product taken with vacuum states $|0\rangle_-$ (as $t\to-\infty$) and $|0_+\rangle$ (as $t\to\infty$), in an external current $J_\mu(x)$.

(ii) Suppose we set $J_\mu = 0$; then we have (cf. (71) and (68)), the result that:

$$Z_{QED}[0] = \mathcal{N}\int \mathcal{D}A_\mu\, e^{iS_0[A_\mu]}\,\tilde{Z}_{Dirac}[A_\mu] \qquad (76)$$

where

$$\left.\begin{aligned} Z_{Dirac}[\bar{A}_\mu] &= \det\left|1 + e\gamma^\mu A_\mu(x)\, G_0(x,x')\right| \\ &\equiv \exp\left\{ \mathrm{Tr}\,\ln\left(\frac{G_0(x,x')}{\mathcal{G}(x,x'|\bar{A}_\mu)}\right)\right\} \end{aligned}\right\} \qquad (77)$$

Now from (i) it is clear that we can interpret $Z_{Dirac}[A_\mu]$ as the amplitude that between $t=-\infty$ and $t=\infty$, the QED vacuum, under the influence of a fixed field $\bar{A}_\mu(x)$, will stay in the vacuum state if it started there at time $t=-\infty$. But we can also write

$$\tilde{Z}_{Dirac}[\bar{A}_\mu] = \lim_{t\to\infty}\int d^3\underline{r}\,\langle 0|\underline{r},t\rangle\rangle\,\mathcal{G}(\underline{r},t;0_-|\bar{A}_\mu) \qquad (78)$$

where $\mathcal{G}(x;0|\bar{A}_\mu)$ is the amplitude to start in the vacuum at $t=-\infty$, and end up at $x=(\underline{r},t)$ at time t. We can now integrate up (75), dividing by

$|\mathcal{G}(x,x|A^\mu)|^2$, to get.

$$\left.\begin{aligned} |\mathcal{Z}_{Dirac}[\bar{A}^\mu]|^2 &\equiv |\langle 0_+|\hat{T}\{e^{-\frac{1}{2}e\int d^4x\, J^\mu(x).\bar{A}_\mu(x)}\}|0_-\rangle|^2 \\ &= e^{-\int d^4x\,\Gamma(x)} \end{aligned}\right\} \quad (79)$$

where, in (78), $J^\mu(x)$ is the current from the Dirac fermion (the external current coupling to $A_\mu(x)$ is zero).

Thus, if we can evaluate the determinant in (71), we will find $\Gamma(x)$; we just have

$$\Gamma_{\bar{A}} = \int d^4x\, \Gamma_{\bar{A}}(x) = Tr\ ln\left(\frac{\mathcal{G}(x,x'|\bar{A}_\mu)}{\mathcal{G}_0(x,x')}\right) \quad (80)$$

for some given field configuration $\bar{A}_\mu(x)$.

There are some nice tricks introduced by Schwinger to do the trace; note that $Tr \equiv \int d^4x\,\delta(x-x')\,tr$, where tr traces over internal indices for the Dirac fermions. First we use the same trick as used between eqtns. (71) and (72) to write

$$Tr\ ln\left|\frac{i\gamma^\mu k_\mu - m}{i\gamma^\mu(k_\mu - eA_\mu) - m + i\delta}\right| = Tr\ ln\left|\frac{-i\gamma^\mu k_\mu - m}{-i\gamma^\mu(k_\mu - e\bar{A}_\mu) - m + i\delta}\right| \quad (81)$$

so that we have

$$\left.\begin{aligned} \Gamma_{\bar{A}} &= ln\ \mathcal{Z}_{Dirac}[\bar{A}^\mu] = Tr\ ln\left|\frac{(\gamma^\mu k_\mu)^2 - m^2}{[\gamma^\mu(k_\mu - e\bar{A}_\mu)]^2 - m^2 - i\delta}\right| \\ &= Tr\ ln\left|\frac{k^2 - m^2}{(k-eA)^2 - m^2 + \frac{1}{2}eS_{\mu\nu}\bar{F}^{\mu\nu} + i\delta}\right| \end{aligned}\right\} \quad (82)$$

where by $(k-eA)^2$ we mean $(k^\mu - eA^\mu)(k_\mu - eA_\mu)$, and where

$$S_{\mu\nu} = \frac{i}{2}[\gamma_\mu, \gamma_\nu] \qquad \bar{F}_{\mu\nu} = [\partial_\mu \bar{A}_\nu - \partial_\nu \bar{A}_\mu) \quad (83)$$

as usual.

The next Schwinger trick is to introduce a proper time representation; noting that

$$ln\left|\frac{a}{b}\right| = -\int_0^\infty \frac{ds}{s}\left[e^{isa} - e^{isb}\right] \quad (84)$$

we take the real part of the log, and write

$$\Gamma_{\bar{A}}^{M} = \mathrm{Re}\int d^4x \; \tfrac{1}{2}\,\mathrm{tr}\int_0^\infty \frac{ds}{s}\,\langle x|\, e^{is[(k-eA)^2-m^2+e/2\,\underline{S}\cdot\underline{F}]} - e^{is[k^2-m^2]}\,|x\rangle \qquad (85)$$

where $\underline{S}\cdot\underline{F} \equiv S_{\mu\nu}\bar{F}^{\mu\nu}(x)$. We notice a few things here, viz.,

(i) The expression inside the integral $\int d^4x$ is just what we want, viz., the rate of pair production per unit spacetime volume (ie., the rate per unit time in a volume of $1\ m^3$).

(ii) If we wish, we can rewrite (82) in a number of ways; writing

$$e^{-\Gamma_{\bar{A}}^{M}} = e^{\mathrm{Tr}\ln[G_0^{-1}G[\bar{A}]]} \equiv e^{\mathrm{Tr}\ln K_{\bar{A}}} = \det|\hat{K}_{\bar{A}}| \qquad (86)$$

we now use the heat kernel / zeta-function representation of the determinant of the operator to write

$$e^{-\Gamma_{\bar{A}}^{M}} = e^{-\zeta_{\bar{A}}} = e^{\ln\prod_n \lambda_n} = e^{\sum_n \ln\lambda_n} \qquad (87)$$

where the $\{\lambda_n\}$ are the eigenvalues of the operator $\hat{K}_A$. Using the same trick (84), we can then write this as

$$\Gamma_{\bar{A}}^{M} = -\sum_n \int_0^\infty \frac{ds}{s}\, e^{i\lambda_n s} \qquad (88)$$

so that we should view the exponentials in (85) as sums over the exponentials of eigenvalues of $\hat{K}_A$.

(iii) The only effect of the fermionic spin in (85) is in the term $e/2\,\underline{S}\cdot\underline{F}$ in the first exponential. Thus if we wished to do this theory for spinless fermions (ie., for scalar electrodynamics), we would get, in place of (85), the result

$$\Gamma_{\bar{A}}^{\mathrm{scalar}} = \mathrm{Re}\int d^4x\; 2\int_0^\infty \frac{ds}{s}\,\langle x|\, e^{is[(k-eA)^2-m^2]} - e^{is[k^2-m^2]}\,|x\rangle \quad (\text{spinless}) \qquad (89)$$

where the extra factor of 4 comes from the sum over Dirac indices.

The expressions above were for an arbitrary configuration $\bar{A}_\mu(x)$, where $\bar{E}(x)$ and $\bar{B}(x)$ vary arbitrarily with spacetime position. Let's now specialize to the case we are interested in, viz., that of a constant electric field E_0; for definiteness we take this to be along the $\hat{z}$-axis, so that the operator in the exponential

in (89) reads (here $k_\perp k^\perp \equiv (k_x k^x + k_y k^y)$):

$$\left.\begin{aligned}(k-eA)^2 - m^2 &\equiv (k_\mu - eA_\mu)(k^\mu - eA^\mu) - m^2 \\ &= k_0 k^0 - k_\perp k^\perp - (k_3 + eE_0 t)(k^3 + eE_0 t)\end{aligned}\right\} \qquad (90)$$

where have used $\bar{A}(x) = (0; 0, 0, -E_0 t)$. We also have the result that

$$\mathrm{tr}\, e^{\frac{i}{2} es S_{\mu\nu} F^{\mu\nu}} = 4 \cosh seE_0 \qquad (91)$$

for the spin trace, as can be verified by direct substitution.

To evaluate (85) or (89), we will first do the trace operations, in order to reduce the problem to an integral over proper time s; we will then do the proper time integral.

The trace over the field-independent term in (85) and (89) is simple; we find that

$$\mathrm{Re}\, \tfrac{i}{2}\, \mathrm{tr} \int d^4x\, \langle x | e^{is(k^2 - m^2)} | x \rangle = \mathrm{Re}\left[\frac{-i}{4\pi^2} \frac{1}{s^3} e^{-ism^2} \right] \qquad (92)$$

by simply converting from $\int d^4x$ to $\sum_k \equiv \int d^4k/(2\pi)^4$. * Notice that the term in m^2, since it is independent of k_μ, just comes out of the integration.

The trace over the field-dependent part is more complicated. We already have the contribution from the spin part of this in (91); we now need to do the momentum trace. To do this we make a few changes of variable. First we write

$$\left.\begin{aligned} x_0 \equiv t \quad &\rightarrow \quad (\tilde{x}_0 - k_3/eE_0) \equiv (\tau - k_z/eE_0) \\ \text{so that} \quad k_0^2 - k_\perp^2 - (k_3 + eE_0 x_0)^2 \quad &\rightarrow \quad k_0^2 - k_\perp^2 - (eE_0 \tilde{x}_0)^2 \end{aligned}\right\} \qquad (93)$$

We then write

$$\left.\begin{aligned} k_0 &\equiv K_0/\sqrt{2} \\ \tilde{x}_0 &\equiv q_0/\sqrt{2} \\ ieE_0 &\equiv \omega_0 \end{aligned}\right\} \qquad (94)$$

where K_0 and q_0 represent the momentum and position of an oscillator with imaginary frequency ω_0. We can then rewrite the trace in a much more obvious way,

---

* It is simplest to actually switch to $\int dx_0 \int \frac{d^3k}{(2\pi)^3}$, keeping the time integral, because of the dependence on time x_0 in the integrand. Note we are integrating over a unit spacetime volume, which can be done by integrating to $\pm\infty$, and then dividing by the (infinite) volume.

in the form

$$\left.\begin{aligned} \mathcal{I}(s;E_0,m) &= \mathrm{Re}\ \mathrm{tr}\int d^4x\,\tfrac{1}{2}\langle x|e^{is[(k-eA)^2-m^2]}|x\rangle \\ &\equiv \tfrac{1}{2}\,\mathrm{Re}\ \mathrm{Tr}\left\{e^{is\left[\frac{1}{2}(K_0^2+\omega_0^2 q_0^2)-k_\perp^2-m^2\right]}\right\} \\ &= \mathrm{tr}\,\tfrac{1}{2}\int\frac{dk_z}{2\pi}\int\frac{d^2k_\perp}{(2\pi)^2}\,e^{-isk_\perp^2}\sum_{n=0}^{\infty}e^{is(n+\frac{1}{2})\omega_0}\,e^{-ism^2} \\ &= \mathrm{tr}\,\tfrac{i}{4}\int\frac{dk_z}{2\pi}\int\frac{d^2k_\perp}{(2\pi)^2}\,e^{-is(k_\perp^2+m^2)}\left(\frac{i}{\sin(\frac{1}{2}s\omega_0)}\right) \end{aligned}\right\} \quad (95)$$

where we use the well-known result for the spectrum of an oscillator in ordinary QM.

To do the integration over the 3-momentum, we need to note that the change of variable in (93) now implies that the k_z integral is linked to the time integral; if the time integral is between times t_1 and t_2, the k_z integral is between 0 and $eE_0(t_2-t_1)$. Taking this time interval to be unity, so that $\int dk_z \to$

$$\mathcal{I}(s;E_0,m) = \frac{1}{16\pi^2}eE_0\ \mathrm{Re}\left\{\frac{-i}{s}e^{-ism^2}\frac{i}{\sinh(eE_0 s)}\right\} \quad (96)$$

We now assemble all the bits of this calculation together, ie., we take the results from (91) and (96) together, and then subtract off (92), to get

$$\Gamma_{\bar{A}} = -\frac{1}{4\pi^2}\,\mathrm{Re}\left[i\int_0^\infty\frac{ds}{s^2}\left(eE_0\coth(eE_0 s)-\frac{1}{s}\right)e^{-ism^2}\right] \quad (97)$$

for Dirac fermions; for scalar QED one gets a result in which, after the trace tr, one gets

$$\Gamma_{\bar{A}}^{\text{scalar}} = \frac{1}{8\pi^2}\,\mathrm{Re}\left[i\int_0^\infty\frac{ds}{s^2}\left(\frac{eE_0}{\sinh(eE_0 s)}-\frac{1}{s}\right)e^{-ism^2}\right] \quad (98)$$

These integrals can be done using contour integration – we notice that the integrands are even functions of s (we replace $\mathrm{Re}\{ie^{-ism^2}\}$ by $\sin(sm^2)$), and so we can extend the integration range to $-\infty < s < \infty$, and then do the integration by extending the integration contour to circle the poles of the integrand.

The final results obtained come as a result of summing these

pole contributions, and one finds that

$$\Gamma_{\bar{A}} = \frac{e^2 E_0^2}{8\pi^2} \sum_{n=1}^{\infty} \frac{1}{n^2} e^{-n\pi m^2/eE_0} \qquad \text{(Dirac fermions)} \qquad (99)$$

and

$$\Gamma_{\bar{A}}^{\text{scalar}} = \frac{e^2 E_0^2}{16\pi^2} \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n^2} e^{-n\pi m^2/eE_0} \qquad \text{(scalar electrodynamics)} \qquad (100)$$

which is the well-known result obtained by Schwinger, and in related form by Heisenberg & Euler, and by Weisskopf *. Notice that these results are similar to that in (74), except that we now have a sum over multiple contributions, with only the 1st term in each series in (99) and (100) corresponding to the tunneling result. The precise interpretation of (99) and (100), and the comparison with (74), is still a matter of debate **, but the key difference between the Schwinger expression and a simple tunneling calculation is that Schwinger's result describes a "vacuum persistence probability", in which one looks at the ongoing effect of a static field over a long period of time, whereas the tunneling probability refers simply to the production of a single pair. One can then argue that over a long period of time, correlations will build up between the different pairs produced at different times by the field, and that these are built into (99) and (100).

If one now asks how relevant this calculation is to the real world, one has to confront two key issues:

(i) The electric field required to see significant effects in QED is colossal; for the exponent in (74) to be $\sim O(1)$, a field $E_0 \sim 10^{18}\ Vm^{-1}$ is needed. Although very intense lasers can produce fields approaching this (and they also exist close to heavy nuclei, causing their instability) it is much better to go to condensed matter analogues. One that is currently favoured is graphene - this is discussed in the paper of Cohen & McGady referenced in the footnote. **

(ii) The electron-positron pairs produced will react back on the vacuum, changing the probability for future pairs to be produced - one ends up with a plasma of electron-positron pairs, and a transport eqtn is required. For a recent discussion, see the paper of Gelis and Tanji in the footnote. **

---

* See J. Schwinger, Phys. Rev. 82, 664 (1951) & the papers of Heisenberg & Euler, & of Weisskopf, deriving less explicit expressions for weak fields, and one in German (they date to 1936).

** See, eg., T.D. Cohen, DA McGady, PR D78, 036008 (2008); F. Gelis, N Tanji, Prog. Part. Nucl. Phys 87, 1 (2016).

Finally, let us stand back a little from this calculation, and ask how we can frame it in terms of a field-dependent action for the Dirac fermions. Let us go back to eqtn (96), which writes the entire QED generating functional as a functional integral over all possible configurations of $A_\mu(x)$, with a weighting factor $Z_{Dirac}[A_\mu]$. It thus follows, by comparing eqtns. (77), (79), and (94), that we can write

$$\left.\begin{aligned} \bar{Z}_{Dirac}[\bar{A}_\mu] &= \exp\left\{-\mathrm{Tr}\,\ln\left(\frac{\mathcal{G}(x,x'|\bar{A}_\mu)}{G_0(x,x')}\right)\right\} \\ &= \mathcal{N}\, e^{\frac{1}{\hbar}\left(S_{Dirac}[\bar{\psi},\psi] + \delta S[\bar{\psi},\psi;\bar{A}_\mu]\right)} \end{aligned}\right\} \qquad (101)$$

where $\delta S[\bar{\psi},\psi;\bar{A}_\mu]$ is the change in the free Dirac action caused by the external field $\bar{A}_\mu(x)$. From (94), we see that in a constant electric field, the change δS is given by

$$\delta S[E_0] = \frac{1}{8\pi^2}\int d^4x \int \frac{ds}{s^2}\left(eE_0 \coth(eE_0 s) - \tfrac{1}{s}\right) e^{-ism^2} \qquad (102)$$

where we note that (94) gives twice the imaginary part of this (because $\Gamma_A = -2\,\mathrm{Im}\,\delta S$, the factor of 2 coming because Γ_A comes from the square of dG/dt — see eqtn. (95)).

One can calculate the change in S coming from a more general field configuration. There are various ways to write this; for example

$$\delta S[\bar{A}_\mu] = \frac{1}{8\pi^2}\int d^4x \int \frac{ds}{s}\left(e^2 \coth(eas)\cot(ebs)\,\underline{E}\cdot\underline{B} - \tfrac{1}{s^2}\right) e^{-ism^2} \qquad (103)$$

where a and b are defined as

$$\left.\begin{aligned} a^2 &= \tfrac{1}{4}\left\{\left[(F_{\mu\nu}F^{\mu\nu})^2 + (F_{\mu\nu}\tilde{F}^{\mu\nu})^2\right]^{1/2} - F_{\mu\nu}F^{\mu\nu}\right\} \\ b^2 &= \tfrac{1}{4}\left\{\left[(F_{\mu\nu}F^{\mu\nu})^2 + (F_{\mu\nu}\tilde{F}^{\mu\nu})^2\right]^{1/2} - F_{\mu\nu}F^{\mu\nu}\right\} \end{aligned}\right\} \qquad (104)$$

so that

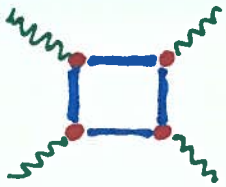
$$\left.\begin{aligned} a^2 - b^2 &= -\tfrac{1}{2}F_{\mu\nu}F^{\mu\nu} = E^2 - B^2 \\ ab &= -\tfrac{1}{4}F_{\mu\nu}\tilde{F}^{\mu\nu} = \underline{E}\cdot\underline{B} \end{aligned}\right\} \qquad (105)$$

and $\tilde{F}_{\mu\nu}(x)$ is the dual EM tensor. Now this correction δS has interactions between photons in it; expanding in powers of e, one gets

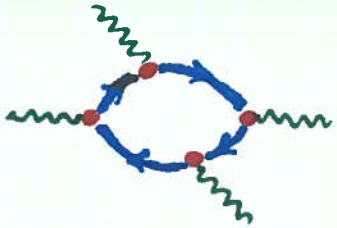
$$\delta S[\bar{A}_{\mu(x)}] = \frac{1}{360\pi^2}\frac{e^4}{m^4}\int d^4x\left[(E^2(x) - B^2(x)) + 7(\underline{E}(x)\cdot\underline{B}(x)\right] + \cdots \qquad (106)$$

which, diagrammatically, corresponds to a term in the action of the form shown in the figure.

$\delta S[\bar{A}_\mu]$ =



≡



Physically we can think of this "light-light" scattering diagram as being a non-linear correction to the strictly linear Maxwell dynamics of a vacuum. The point is that 2 photons can now interact via a quantum fluctuation of a vacuum – the quantum fluctuation is just an electron loop, or electron-positron pair. The 2 ways of drawing this shown in the figure correspond to 2 different ways of thinking about it physically, viz.,

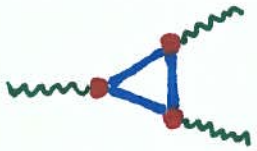
(i) We can think of 2 incoming photons interacting with a spontaneously created electron loop – the fermion "box" shown above – and then being re-radiated at a later point. Thus we have generated a photon-photon scattering term.

(ii) We can think of a photon creating an electron-position pair, which then interacts with 2 photons before recombining with the emission of a photon.

Whichever way one decides to think about this diagram, we see that because of it (and higher-order contributions), the ordinary QED vacuum has acquired dielectric properties – its electric and magnetic susceptibilities are no longer those of the classical EM vacuum. To put it another way – a very intense photon beam will polarize the vacuum, modifying its properties in a way that changes its dielectric properties – its electric & magnetic permeabilities – in a way that modifies the propagation of light itself.*

Note that the lowest contribution to the correction $\delta S[\bar{A}_\mu]$ is of 2nd order in α_{QED}; there is no linear term, ie., no diagram of the form shown in the figure (the "triangle graph"). Such a

Triangle Graph :



gives zero. However in section 6.4 we will find that for chiral QED (when the fermion mass $m \to 0$), there is "chiral anomaly", and this diagram is non-zero

---

For a recent review of this material, see G.V. Dunne, /arXiv 1202.1557

(ii) THE CASIMIR EFFECT

Another very straightforward way that we can imagine perturbing the vacuum is by allowing it to be constrained geometrically. The first discussion of this were given by Casimir & by Casimir & Polder in 1948, who were initially interested in generalizing the theory of van der Waals interactions to incorporate relativistic retardation.* However at this point Casimir, apparently after a conversation with N. Bohr, realized that one could think of the calculation in a rather different way, as a modification of the vacuum state (and its energy) by the interacting objects. To demonstrate this he showed how to calculate the interaction between 2 neutral conducting plates as a pressure caused by the exclusion of certain EM vacuum modes from the region between the plates.

Since this early work the topic has become of great interest in many different fields of physics. Amongst the key developments are:

(i) a fundamental work by E.M. Lifshitz in 1956, which treated the problem using macroscopic electrodynamic theory — later extended by Lifshitz et al. to a general QFT treatment — for the general problem of 2 dielectrics interacting via a 3rd intervening dielectric. The theory was framed in terms of interacting currents and dipoles, and had the great advantage of giving detailed predictions for a huge variety of realistic physical situations.**

(ii) A broad array of theoretical techniques was developed to analyze the Casimir effects in a large variety of physical settings. Most notable amongst these are zeta function / heat kernel techniques (these techiques being required to regularize the divergent integrals) and multiple-scattering/diagrammatic techniques (to deal with complex geometries).***

(iii) An increasingly sophisticated battery of experimental techniques have been brought to bear on the problem, allowing tests of the theory in a whole variety of systems. This has now been done for different metals, semiconductors, & insulators, in geometries ranging from plates, single plates with spheres, hemispheres, etc., to cylinders (ranging down to nanotubes), and to torque measurements on different systems interacting with plates. Many different biological phenomena, involving membranes or long chain molecules, also involve quite strong Casimir / van der Waals forces. In such work, and theoretical attempts to

---

* HBG Casimir, D. Polder, Phys. Rev. 73, 360 (1948); HBG Casimir, Proc. Kon. Ned. Akad. Wet. 51, 793 (1948)

** EM Lifshitz, JETP 2, 73 (1956); IE Dzyaloshinskii, EM Lifshitz, LP Pitaevskii, Adv. Phys. 10, 165 (1961)

*** For heat kernel techniques, see the Appendix on regularization techniques. For the application of multiple scattering techniques to the Casimir effect, see R. Balian, B. Duplantier, Ann. Phys 104, 300 (1977), and 112, 165 (1978); & MF Maghrebi et al., Phys. Rev. D83, 045004 (2011), and D87, 025016 (2013), and refs therein.

deal with the experiments, the effects of temperature and temperature fluctuations turn out to be crucial. *

(iv) The generalization of these ideas to various relativistic QFT's, and the application to cases of confined geometries (eg., the atomic nucleus). Most important, the treatment of QFT's in curved spacetimes, which raises many conceptual & mathematical questions. The use of idealized models, involving mirrors or sets of mirrors, accelerating detectors, and spacetimes of reduced dimensionality, has been crucial in advancing the theory. Well-known results, such as the existence of Hawking radiation, have followed from this work, which also plays a key role in modern cosmological theory. **

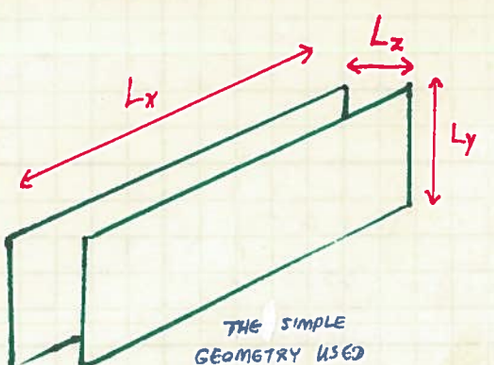
More generally, it is clear that any attempt to confine some field must come through its interaction with some other field, which we typically treat as a "background field". Even the spacetime field falls into this category. Depending on the configuration of the background field, this can lead to a large variety of effects, including the Schwinger effect just discussed, to Hawking radiation, and various condensed matter analogues of these, as well as forces like the Casimir force.

In what follows we only scratch the surface of this very large topic.

- DERIVATION VIA MODE-COUNTING :

The simplest & most obvious way to address the Casimir energy is via the counting of modes inside & outside some closed or semi-closed container. The discussion given here is similar in spirit (although very different in detail) to that done by Casimir in 1948. The method is very limited, however, since it only applies to containers with special boundary conditions on the fields – typically these only apply to perfect conductors, when we deal with EM fields.

We consider the geometry shown in the picture; 2 conducting plates are separated by L_z, and they are assumed to extend in infinite distance in the xy-direction (ie, the container or box is assumed to have dimensions L_x, L_y, and L_z, with $L_x, L_y \to \infty$).



THE SIMPLE GEOMETRY USED BY CASIMIR: 2 PLATES, SEPARATED BY A DISTANCE L_z, EXTENDING TO $\pm\infty$ IN THE XY-PLANE.

The Casimir calculation begins by ignoring QED particle-hole pair fluctuations, and calculates the difference between the vacuum energy density integrated

* See, eg., GL Klimchitskaya et al., Rev. Mod. Phys. 81, 1827 (2009); I Brevik et al., NJP 8, 236 (2006).

**

over the volume between the plates, and the energy density in the absence of the plates. We then write (with $L_\perp^2 = L_x L_y$):

$$\bar{E}_{Cas}(L_z) = \lim_{L_\perp \to \infty} \frac{1}{L_\perp^2 L_z} \int_V d^3r \, E_{Cas}(\underline{r}) = \lim_{L_\perp \to \infty} \frac{1}{L_\perp^2 L_z} \int_V d^3r \left[E_{pl}(\underline{r}) - E_{Vac}(\underline{r}) \right] \tag{107}$$

for the MEAN Casimir energy density, with $\int_V d^3r \equiv \int_{-L_x/2}^{L_x/2} dx \int_{-L_y/2}^{L_y/2} dy \int_{-L_z/2}^{L_z/2} dz$ being the volume integral between the plates.

The obvious way to calculate this is by converting to a sum over modes in $\underline{k}$-space; to do this we need to know what are the allowed photon modes between the plates in this case. We recall that there are 2 polarizations for each wave-vector $\underline{k}$, for EM modes; and that for a perfect conducting surface, the magnetic field is continuous across the boundary, whereas the electric field drops to zero at the surface (we assume no surface charge). We can then discuss the whole problem in terms of 2 scalar modes, representing the 2 (transverse) photon modes, having the form

$$\left.\begin{aligned} &(i)\quad \phi_{\underline{k}}(\underline{r}) = \cos\left(\frac{\pi n z}{L_z}\right) e^{i\underline{k}_\perp \cdot \underline{r}} e^{-i\omega_n(k_\perp)t} \qquad \left(\text{Dirichlet B.C.: } \phi_{\underline{k}}\left(\pm \tfrac{L_z}{2}\right) = 0\right) \\ &(ii)\quad \phi_{\underline{k}}(\underline{r}) = \sin\left(\frac{\pi n z}{L_z}\right) e^{i\underline{k}_\perp \cdot \underline{r}} e^{-i\omega_n(k_\perp)t} \qquad \left(\text{Neumann B.C.: } \partial_z \phi_{\underline{k}}\left(\pm L_z/2\right) = 0\right) \end{aligned}\right\} \tag{108}$$

where the eigenenergies are the same for these 2 modes, ie.,

$$\hbar\omega_n(k_\perp) = \frac{\hbar c}{2}\left[k_\perp^2 + \left(\frac{\pi n}{L_z}\right)^2\right]^{1/2} \tag{109}$$

There is one mode excluded from the above set - this is the $k_z = 0$ Neumann mode, corresponding to a constant $\underline{B}$-field, which also satisfies the boundary condition $\partial_z \phi_{\underline{k}}(\pm L_z/2) = 0$. Thus, adding all the energies together, we get

$$\bar{E}_{Cas} = \lim_{L_\perp^2 \to \infty} \frac{1}{L_\perp^2 L_z} \frac{\hbar c}{2} \int \frac{dk_\perp^2}{(2\pi)^2} L_\perp^2 \left\{ L_z \left[\left(\overset{\text{Neumann}}{\sum_{n=0}^{\infty}} + \overset{\text{Dirichlet}}{\sum_{n=1}^{\infty}} \right) \left(k_\perp^2 + \left(\frac{n\pi}{L_z}\right)^2 \right)^{1/2} \right] - 2L_z \int_{-\infty}^{\infty} \frac{dk_z}{2\pi} \left(k_\perp^2 + k_z^2\right)^{1/2} \right\} \tag{110}$$

where the sum is divided into Neumann & Dirichlet B.C. terms, and the last term is the vacuum contribution in the absence of the plates, inside

the region where the plates were.

There are many ways to deal with the sums in (110). The key problem here is that the sums and integrals diverge, and need to be regulated. This regulation, ie., cutting off of the high $|k|$ contributions, corresponds physically to the fact that at sufficiently high frequency, any conductor (and indeed any real physical system) becomes transparent to EM waves - the dielectric & susceptibility functions $\epsilon(\omega)$ and $\mu(\omega)$ tend to their vacuum values at high frequencies. We discuss this point below.

However - and this is a key point, that needs to be verified - we can try arguing that the physical answer should not depend on the very high energy behaviour. If this argument is correct, then the result we calculate should not depend on the method of regularization. In what follows we will verify that this is the case, for this particular problem.

We will derive our final answer here in 3 different ways (later on we will rederive it in other, more general ways), which illustrate different parts of the physics. The first 2 ways are modern, the last is essentially the method used by Casimir

METHOD 1 : ζ-FN REGULARIZATION :

To make this quick and simple, lets rewrite (110) in the form

$$\bar{E}_{Cas} = \frac{\hbar}{2}\int\frac{d^2k_\perp}{(2\pi)^2}\,2\sum_{n=1}^{\infty}\omega_n(k_\perp) - \frac{\hbar}{2}\int\frac{d^2k_\perp}{(2\pi)^2}\left[\left(\int_{-\infty}^{\infty}\frac{dk_z}{2\pi}\,2\,\omega_{k_\perp}(k_z)\right) - \omega_{k_\perp}(0)\right] \quad (111)$$

where $\omega_n(k_\perp)$ is given by (109), and $\omega_{k_\perp}(k_z)$ is just the vacuum mode energy, ie.,

$$\omega_{k_\perp}(k_z) = \frac{\hbar c}{2}\left[k_\perp^2 + k_z^2\right]^{1/2} \quad (112)$$

We now observe that only the 1st term in (111) actually depends on L_z (via the dependence of $\omega_n(k_\perp)$ on L_z). Both terms are of course infinite - the first refers to the energy when we have plates, the second to when we don't (or equivalently, to when we let $L_z \to \infty$).

If we are only interested in the Casimir energy, we can try to isolate out the L_z-dependent part - the easiest way is to use ζ-function regularization. To do this, we write

$$\Delta\bar{E}_{Cas} = \hbar\int\frac{d^2k_\perp}{(2\pi)^2}\sum_{n=1}^{\infty}\omega_n(k_\perp) = \lim_{s\to -1}\hbar\int\frac{d^2k_\perp}{(2\pi)^2}\sum_{n=1}^{\infty}|\omega_n(k_\perp)|^{-s} \quad (113)$$

ie., we have replaced $\omega_n(k_\perp)$ in the sum by $\omega_n(k_\perp)\,|\omega_n(k_\perp)|^{-(1+s)}$. The

idea is to calculate this more general function, and then analytically continue back to $s = -1$. We have, explicitly, that

$$\begin{aligned}
\Delta\bar{E}_{Cas}(s) &= \hbar \int \frac{d^2k_\perp}{(2\pi)^2} \sum_{n=1}^{\infty} |\omega_n(k_\perp)|^{-s} \\
&= \frac{\hbar c^{-s}}{2\pi} \sum_{n=1}^{\infty} \int_0^{\infty} k_\perp \, dk_\perp \left| k_\perp^2 + \left(\frac{\pi n}{L_z}\right)^2 \right|^{-s/2} \\
&= \frac{\hbar c^{-s} \pi^{1-s}}{4} \frac{\Gamma(1-s/2)}{\Gamma(s/2)} \frac{1}{L_z^{2-s}} \sum_{n=1}^{\infty} |n|^{2-s} \\
&= \frac{\hbar}{2} \frac{\pi^{1-s}}{s-2} \frac{c^{-s}}{L_z^{2-s}} \zeta_R(2-s)
\end{aligned} \tag{114}$$

where $\zeta_R(z)$ is the Riemann ζ-function, defined by $\zeta_R(z) = \sum_{n=1}^{\infty} n^{-z}$. The problem arises when we take $s \to -1$; we then get

$$\Delta\bar{E}_{Cas} = -\hbar \frac{\pi^2 c}{6} \frac{1}{L_z^3} \zeta_R(-3) \tag{115}$$

and of course $\zeta_R(-3)$ is divergent. At this point one makes the analytic continuation by what looks at first glance like a trick – we use the "reflection identity" for the ζ_R-function, viz.,

$$\zeta_R(1-z) = \pi^{1-z} \frac{\Gamma(z/2)}{\Gamma(1-z/2)} \zeta_R(z) \tag{116}$$

to define $\zeta_R(-3)$ as

$$\begin{aligned}
\zeta_R(-3) &= \pi^{-7/2} \frac{\Gamma(5/2)}{\Gamma(-3/2)} \zeta_R(4) \\
&= 1/120
\end{aligned} \tag{117}$$

so that we finally get, by combining (117) and (115), that

$$\Delta\bar{E}_{Cas}(L_z) = -\hbar \frac{\pi^2 c}{720} \frac{1}{L_z^3} \tag{118}$$

The use of the Riemann reflection formula has allowed us to get rid of an infinite contribution – how legitimate this will be discussed below, once we

have looked at the other methods. We note from (118) that we get a negative pressure on the plates, given by

$$\bar{P}(L_z) = -\frac{d}{dL_z}\bar{E}_{Cas}(L_z) = -\hbar\frac{\pi^2 c}{240}\frac{1}{L_z^4} \tag{119}$$

which is often (incorrectly) described as being due to the deficit of modes between the plates, as compared to those outside - this will become clear when we consider other geometries.

METHOD 2 : EXPONENTIAL CUT-OFF : The ζ-fn method just given imposes a power-law cut-off on the sum over modes. If instead we impose an exponential cut-off, ie., we write

$$\bar{E}_{Cas}(L_z) = \lim_{\Lambda_0\to\infty}\frac{\hbar}{2}\int\frac{d^2k_\perp}{(2\pi)^2}\left[\sum_{n=-\infty}^{\infty}\omega_n(k_\perp)\,e^{-\omega_n(k_\perp)/\Lambda_0} - 2L_z\int\frac{dk_z}{2\pi}\,\omega_{k_\perp k_z}\,e^{-\omega_{k_\perp k_z}/\Lambda_0}\right] \tag{120}$$

with $\omega_n(k_\perp)$ and $\omega_{k_\perp k_z}$ given in (109) and (112). Changing variables, so that $x = k_\perp L_z/\pi$, so that we have $d^2k_\perp = (\pi/L_z)^2\,d^2x$, and writing $n = L_z k_z/\pi$, we have

$$\bar{E}_{Cas}(L_z) = \frac{\pi^2\hbar c}{L_z^3}\lim_{\Lambda_0\to\infty}\int d^2x\left\{\sum_{n=0}^{\infty}(x^2+n^2)^{1/2}\,e^{-\omega(x,n)/\Lambda_0} - \int_0^\infty du\left[(n^2+u^2)^{1/2}e^{-\omega(x,n)/\Lambda_0} - \tfrac{1}{2}n\right]\right\} \tag{121}$$

where we have $\omega(x,n) = \frac{\pi c}{L_z}(x^2+n^2)^{1/2}$.

Now let us deal directly with the difference between the sum and integral in this formula. The standard way to do this is using the Euler–MacLaurin formula, discussed in the next method - but here we use a trick introduced by Starobinski et al. in 1976*. This uses the Abel-Plana identity, viz.,

$$\sum_{n=0}^{\infty}F_n - \tfrac{1}{2}F(0) - \int_0^\infty dn\,F(n) = i\int_0^\infty\frac{dn}{e^{2\pi n}-1}\left[F(in)-F(-in)\right] \tag{122}$$

where we consider the function $F(n)$ of a continuous variable n, and F_n is defined for n = integer. We note the connection to the Euler-MacLaurin identity, viz.

$$\sum_{n=1}^{\infty}F_n + \tfrac{1}{2}F(0) - \int_0^\infty dn\,F(n) = -\sum_{n=1}^{\infty}B_{2n}\frac{d^{2n-1}}{dn}F(n)\Big|_{n=0} \tag{123}$$

---

* S.G. Mamayev, V.M. Mostepanenko, A.A. Starobinsky, JETP 43, 823 (1976)

where the B_n are the Bernoulli numbers, defined by

$$\frac{x}{e^x - 1} = \sum_{n=0}^{\infty} B_n \frac{x^n}{n!} \tag{124}$$

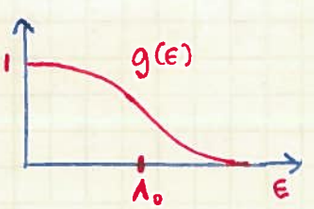
Applying the Abel-Plana result, one gets (letting $\Lambda_0 \to \infty$)

$$\left.\begin{aligned} \bar{E}_{Cas}(L_z) &= -\hbar \frac{\pi^2 c}{L_z^3} \int_0^\infty x\,dx \int_x^\infty dn \frac{(n^2 - x^2)^{1/2}}{e^{2\pi n} - 1} \\ &= -\hbar \frac{\pi^2 c}{3 L_z^3} \frac{1}{(2\pi)^4} \int_0^\infty \frac{x^3\,dx}{e^x - 1} \\ &= -\hbar \frac{\pi^2 c}{720} \frac{1}{L_z^3} \end{aligned}\right\} \tag{125}$$

as before. We notice that in this calculation, the cancellation between the 2 infinite terms has been absorbed directly.

METHOD 3 : GENERIC CUT-OFF : We now come finally to the method employed by Casimir himself to deal with this calculation. Instead of the exponential regulator employed in (121), we now employ a regulator function $g(\epsilon)$, we put the integration in (121) in a form suitable for the Euler-MacLaurin expansion, viz.,

$$\bar{E}_{Cas}(L_z) = \frac{\pi^2 \hbar c}{2 L_z^3} \int_0^\infty x\,dx\; g\left(\frac{\pi}{L_z}(x^2 + n^2)^{1/2}\right) \left[\frac{x}{2} + \left(\sum_{n=1}^{\infty} - \int_0^\infty dn \right) (x^2 + n^2)^{1/2} \right] \tag{126}$$



where we again go back to the form with a sum from $n = 1$ to $n = \infty$ (rather than $\sum_{n=0}^{\infty}$ as in (121)).
If we now define a regularized function

$$f(x, n) = (x^2 + n^2)^{1/2}\, g\left(\frac{\pi}{L_z}(x^2 + n^2)^{1/2}\right) \tag{127}$$

such that not only $g(\epsilon)$, but also $f(\epsilon)$, go to zero for large enough ϵ (in practice, above a UV cutoff Λ_0), then we can write this as

$$\bar{E}_{Cas}(L_z) = \hbar \frac{\pi^2 c}{2 L_z^3} \left\{ \tfrac{1}{2} F(0) + \sum_{n=1}^{\infty} F_n - \int_0^\infty dn\, F(n) \right\} \tag{128}$$

where we define $F(n) = \int_0^\infty x\,dx\, f(x, n)$, and where F_n is as before given

by $F(n = \text{integer}) = F_n$. We may now use the Euler-MacLaurin expansion in (123) directly. We assume that $g(0) = 1$, and also that all its derivatives are zero at $\epsilon = 0$, to simplify things. It then follows that the only non-zero derivative of $F(n)$ is $d^3F(n)/dn^3$, and since $B_4 = -1/30$, we get again

$$\bar{E}_{Cas}(L_z) = \hbar \frac{\pi^2 c}{L_z^3} \frac{B_4}{4!} = -\hbar \frac{\pi^2 c}{720} \frac{1}{L_z^3} \qquad (129)$$

It is remarkable that Casimir, in 1948, was able to see his way through to this result, including the deployment of what we would now call a UV regulator, before the formal apparatus of relativistic QED as a covariant field theory had been constructed, and well before a serious understanding of the QED vacuum had been attained. *

ANYPAD

- DERIVATION VIA DIELECTRIC FLUCTUATIONS :

The discussion up to now has given the impression that one can understand the Casimir effect purely in terms of the vacuum modes - the only function of the boundaries is to restrict these in some way. It therefore comes as a surprise to realize that there is a quite different way of thinking about Casimir energies, as arising from van der Waals interactions between fluctuating dipoles. This was in fact the original motivation of Casimir & Polder - it was only afterwards that Casimir realized the connection to vacuum modes.

What follows is an extremely abbreviated discussion — there have been literally thousands of papers discussing the different ways of calculating interactions between 2 or more dielectric solids, or between solids & liquids, or between liquids, in a wide variety of fields ranging from quantum gravity & relativistic QFT, through solid state physics to biology. Here the main points to be made are (i) that one can think of the Casimir effect as arising from interactions between modes/excitations connected with the interfaces between different media, and that (ii) these interactions are essentially identical to, or a generalization of, the standard van der Waals interactions from atomic physics. In what follows we first recall a few things about van der Waals interactions, and then sketch the derivation of the famous Lifshitz formula for the general interaction between different plane-faced media. This will make clear the main points, which have also been discussed in similar ways by many other authors. * Note also that another closely related way of doing calculations of this kind involving scattering theory (T-matrix, or related methods, sometimes written with complicated diagrammatic rules); in all these cases one is dealing with the field-mediated interactions between dielectric bodies, either through a vacuum

---

* Apart from papers cited earlier, key references are G. Feinberg, J. Sucher, Phys. Rev. A2, 2395 (1970); J. Schwinger, Lett. Math. Phys. 1, 43 (1975); J. Schwinger, LL DeRaad, K.A. Milton, Ann Phys (NY) 115, 1 (1978); KA. Milton, J. Phys. A37, R209 (2004); S.J. Rahi, T. Emig, N. Graham, RL Jaffe, M. Kardar, Phys. Rev. D80, 085021 (2009); and the many refs. therein

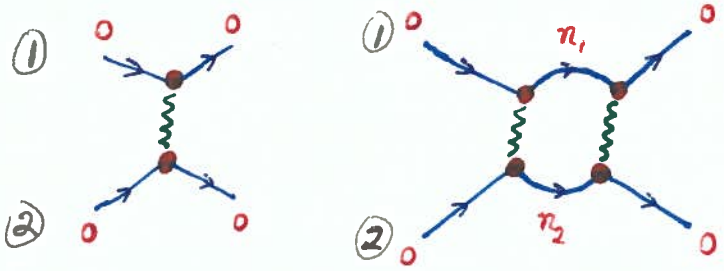
or via some other dielectric medium. Since, as we have already seen, the vacuum itself should be viewed as a dielectric medium, there is no fundamental difference here — the differences between the different dielectrics lie solely in the (virtual) excitations that are polarised, by the fields generated by the relevant charges at the interfaces.

VAN DER WAALS INTERACTIONS

: Lets just begin by quickly recalling what the van der Waals interaction is, and where it comes from. The classic derivation in the non-relativistic limit was given by F. London *, and it gives the force between a pair of polarizable atoms or molecules (or indeed any other very small object) in the regime where the photon-mediated interaction between them can be treated as instantaneous. As London noted, even if neither system possesses an electric dipole moment, it can fluctuate into a state which does — the van der Waals interaction is then an interaction between the virtual states of a pair of systems, in which they do have such a moment. The interaction potential for a pair of dipoles $\underline{d}_1$ and $\underline{d}_2$ is just

$$V(\underline{d}_1;\underline{d}_2;\underline{r}_{12}) = \frac{1}{r_{12}^3}\left[\underline{d}_1\cdot\underline{d}_2 - 3\frac{(\underline{d}_1\cdot\underline{r}_{12})(\underline{d}_2\cdot\underline{r}_{12})}{r_{12}^2}\right] \qquad (130)$$

where $\underline{r}_{12}$ is the vector between the 2 objects. ** Now if $\underline{d}_1$ and $\underline{d}_2$ are zero in the ground state of the systems, we get no direct interaction — but the 2nd-order process does give one.



DIAGRAMS FOR THE SCATTERING OF 2 ATOMS/MOLECULES VIA PHOTONS: THE IN & OUT STATES ARE VACUUM STATES

The diagrams show these 2 processes — note that the spin & 4-momenta of the objects are irrelevant here, and we are only concerned with internal orbital states of these systems, which we label 1 and 2.

The 2nd-order interaction is just given by standard QM as

$$\Delta E = \sum_{n_1 n_2} \frac{|V_{0n_1}|^2 |V_{0n_2}|^2}{E^{(1)}_{n_1} + E^{(2)}_{n_2} - E^{(1)}_0 - E^{(2)}_0} \qquad (131)$$

---

* F. London, Trans. Faraday Soc. 33, 8 (1937)

** The same physics occurs with virtual fluctuations involving magnetic moments if we include spin, via the magnetic dipolar interaction $V(\underline{m}_1,\underline{m}_2;\underline{r}_{12}) = -\frac{1}{r_{12}^3}\left[\underline{m}_1\cdot\underline{m}_2 - 3\frac{(\underline{m}_1\cdot\underline{r}_{12})(\underline{m}_2\cdot\underline{r}_{12})}{r_{12}^2}\right]$

and if one works this out for transitions mediated by $V(d_1, d_2; r_{12})$ between the ground states of an atom in an S-state, and the excited p-orbital states, one gets London's answer, viz.,

$$\Delta E = -\frac{2}{3}\frac{1}{r_{12}^6}\sum_{n_1 n_2}\frac{|d_{0n_1}^{(1)}|^2\,|d_{0n_2}^{(2)}|^2}{\epsilon_0^{(1)}+\epsilon_0^{(2)}-\epsilon_{n_1}-\epsilon_{n_2}} \qquad (132)$$

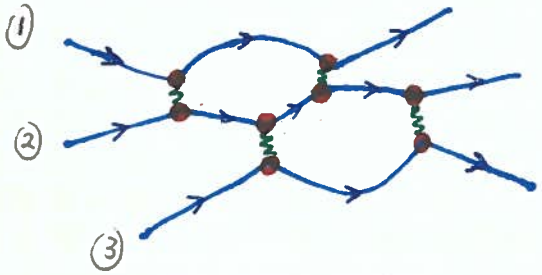
where $d_{0n_j}^{j} \equiv \langle 0|\underline{d}^{j}|n_j\rangle$ is the matrix element between the ground state and the excited p-states $|n_j\rangle$, with $j=1,2$ for the 2 different atoms.

As Casimir & Polder pointed out, at long ranges, retardation changes the $1/r^6$ form to a $1/r^7$ form. To show this one must use the full photon propagator in the intermediate lines. The calculation is rather messy, but the final answer can be written as

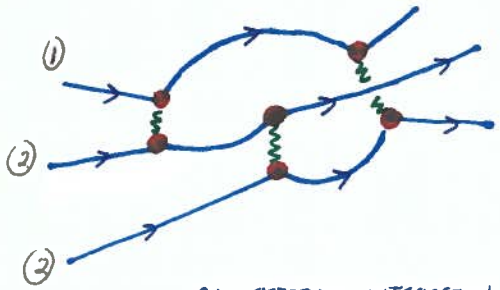
$$\Delta E = -\frac{23}{\pi}\frac{1}{r_{12}^7}\sum_{n_1 n_2}\frac{|d_{0n_1}^{(1)}|^2\,|d_{0n_2}|^2}{\epsilon_0^{(1)}-\epsilon_0^{(2)}-\epsilon_{n_1}-\epsilon_{n_2}} \qquad (133)$$

and the interaction crosses over to this at lengthscales $> \hbar c/(\epsilon_0-\epsilon_n)$, the wavelength of the photons involved in mediating the transition *

One can generalize such calculations to look at interactions between assemblies of different bodies, at which point a key observation needs to be made - the van der Waals interaction is <u>not additive</u>; in other words, we can't derive the interaction between 3 or more systems by simply summing the pairwise interactions between the different systems. A diagrammatic analysis shows why - consider the VdW interaction between 3 systems, as shown in the figure.



AN EFFECTIVE INTERACTION FOR 3 SYSTEMS, 2nd-ORDER FOR SYSTEM ① & ③, 4-TH ORDER FOR SYSTEM ②.



AN EFFECTIVE INTERACTION BETWEEN 3 SYSTEMS, 2nd-ORDER FOR EACH OF THEM.

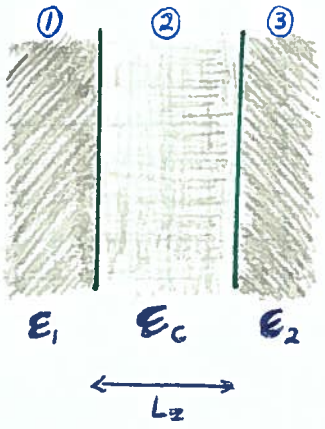
---

* Since one is taking the static limit for both (132) & (133) of the same interaction, why the difference? It is because the static limit involves an $\omega=0$ photon in (132), but to get (133), 2 in general finite-ω photons are involved; the intermediate state is a dynamic dipolar fluctuation.

These "3rd-party" indirect interactions will of course come up any time in QFT for a many-body system - but the difference here is that the photon-mediated interactions are long-range in nature, so that the cumulative effects of a "3rd-party medium" is to change the form of the interaction between any 2 systems, and to make the total effect non-additive (this is obvious in 3d, because we then deal with integrals $\sim \int d^3r/r^3$).

Thus, for macroscopic media, one needs a different approach.

LIFSHITZ THEORY: In a treatment which now stands as something of a landmark in condensed matter physics (and which was done in the middle of work on singularities in general relativity, and work in cosmology), Lifshitz addressed this problem in a very thorough way, for a set of different media having different dielectric properties - the scheme we will discuss here is shown in the figure. One has 2 interfaces, at coordinates $z = 0, L_z$, in which a "cavity medium" of dielectric constant ϵ_c is sandwiched between 2 media having dielectric constants ϵ_1 and ϵ_2. In general these dielectric media have frequency-dependent dielectric function $\epsilon_j(\omega)$; indeed we expect that, no matter what $\epsilon_j(\omega)$ may be when $\omega \to 0$, the media will become transparent for very high ω, ie, we expect



$$\epsilon_j(\omega) \xrightarrow[\omega \gg \Lambda_0]{} \epsilon_0 \qquad (134)$$

for Λ_0 some UV crossover frequency.

We now wish to find the forces in this system - they will be between the 2 interfaces. There are various ways this can be done. Amongst these are

- Scattering theory, as noted above
- Schwinger-style "source theory", where we look at all of the EM modes as sourced by the quantized Dirac currents
- Lifshitz-style analysis, in which the combined matter current/EM field is quantized, & the combined classical + thermal fluctuations are analyzed in a Matsubara formulation.
- A local analysis, in which one calculates the stress-energy tensor $T_{\mu\nu}(\underline{r})$, and then integrates it over the regions of interest.

In what follows I will give a treatment which is a kind of amalgam of all these; but I will emphasize the sum over the mode energies of the new dielectric vacuum, so as to see the connection with what we've already done. I will also introduce yet another trick for doing the sums over these modes.

We start by looking at what the photon modes will look like in these dielectric media. As remarked above, we characterize these media by their dielectric functions $\epsilon_j(\omega)$, and this then defines the propagation of EM waves in

the various media; we have, writing

$$F_{\mu\nu}(x) = F_{\mu\nu}(\underline{r})\, e^{-i\omega t} \tag{135}$$

(the separation of space & time variables necessary because we have lost Lorentz invariance), we can write the eqtn. of motion for photons as

$$\left(c^2\nabla^2 + \omega^2\varepsilon(\omega)\right) F_{\mu\nu}(r) = 0 \tag{136}$$

or in Fourier transformed form as

$$\left.\begin{aligned} k^2 F_{\mu\nu}(k) &= \left(\underline{k}^2 - \varepsilon(\omega)\tfrac{\omega^2}{c^2}\right) F_{\mu\nu}(k) \\ &= \left(k_z^2 + \kappa^2\right) F_{\mu\nu}(k) = 0 \end{aligned}\right\} \tag{137}$$

where $\underline{k}^2 = k_x^2 + k_y^2 + k_z^2$ is the 3-momentum, and $\kappa^2 = \underline{k}_\perp^2 - \varepsilon(\omega)\omega^2/c^2$, with $\underline{k}_\perp^2 = k_x^2 + k_y^2$. Here we separate out k_z from $\underline{k}_\perp$, as before, because the system is not translationally symmetric along $\hat{z}$.

To get the right solutions to Maxwell's eqtns we need boundary conditions @ the interfaces; these are (under the assumption that the magnetic field components of $F_{\mu\nu}(x)$ are irrelevant, so we only deal with $\underline{E}(x)$), as follows:

$$\text{B.C.s:}\quad \left.\begin{aligned} &\varepsilon(\omega)E_z(z),\ \frac{d}{dz}E_z(z) \\ &\underline{E}_\perp(z),\ \frac{d}{dz}\underline{E}_\perp(z) \end{aligned}\right\} \text{ all continuous across interfaces.} \tag{138}$$

We can then write an ansatz for the components of $\underline{E}(\underline{r})$; for the E_z component we write

$$\left.\begin{aligned} E_z(z) &= E_1 e^{K_1 z} && (z<0)\ \text{(region ①)} \\ E_z(z) &= E_+^c e^{K_c z} + E_-^c e^{-K_c z} && (0<z<L_z)\ \text{(region ②)} \\ E_z(z) &= E_2 e^{-K_2 z} && (z>L_z)\ \text{(region ③)} \end{aligned}\right\} \tag{139}$$

and likewise for $\underline{E}_\perp$. Substituting into Maxwell's eqtn. (137), and incorporating the boundary conditions, gives us 2 eqtns for the eigenmodes:

$$\left.\begin{aligned} &\text{(a)}\quad (K_1+K_c)(K_2+K_c)e^{2K_cL_z} - (K_1-K_c)(K_2-K_c) = 0 \\ &\text{(b)}\quad (\varepsilon_cK_1+\varepsilon_1K_c)(\varepsilon_cK_2+\varepsilon_2K_c)e^{2K_cL_z} - (\varepsilon_cK_1-\varepsilon_1K_c)(\varepsilon_cK_2-\varepsilon_2K_c) = 0 \end{aligned}\right\} \tag{140}$$

Now these 2 sets of eqtns determine the allowed set of eigenmodes for the system; but a number of points should be underlined here first:

(i) This is not the complete set of modes for the system - what we have captured here is the set of modes that are localized around the interfaces, under the assumption that K_1, K_2, and K_c are real. There is another whole set of modes, with imaginary values for K, which describe the vacuum modes that extend throughout all space. They make no contribution to the Casimir energy.

(ii) The above eqtns determine a set of energies $\omega_n(k_\perp)$, since they determine allowed values for $K_n^2 = k_\perp^2 - \varepsilon(\omega)\omega_n^2/c^2$. However one cannot satisfy both eqtns (140a) and (140b) simultaneously. The first one, eqtn (140a), can be satisfied if $E_z = 0$, whereas the 2nd one (140b) requires that $E_\perp = 0$. We get 2 different sets of modes for these 2 different conditions.

(iii) Unlike the previous discussion, where the evaluation of the sum over energy eigenvalues could be done directly since we had explicit results for the $\omega_n(k_z)$, here we only have implicit eqtns in (140) for these eigenvalues, which cannot be evaluated in any simple way. However, we only require the sum $\sum_n \omega_n(k_\perp)$; and we shall use a simple trick to get this sum.

Let us now write the Casimir part of the energy for this problem in the form (cf. eqtn.

$$E_c(L_z) = \frac{\hbar}{2}\int \frac{d^2k_\perp}{(2\pi)^2}\left(\sum_{n_\perp}\omega_{n_\perp}(k_\perp) + \sum_{n_z}\omega_{n_z}(k_\perp)\right) \tag{141}$$

in which the sum $\Sigma_{n_\perp}$ refers to modes where $E_\perp = 0$, and $\Sigma_{n_\parallel}$ refers to modes with $E_z = 0$. Since we do not know explicit forms for the $\omega_{n_\perp}(k_\perp)$ and $\omega_{n_z}(k_\perp)$, we use the following trick. Recall the "argument lemma", a corollary of Cauchy's thm. for meromorphic functions, according to which a given meromorphic function $f(z)$ obeys

$$\frac{1}{2\pi i}\oint_C dz\,\frac{f'(z)}{f(z)} \equiv \frac{1}{2\pi i}\oint_C dz\,\ln f(z) = N_0 - N_p \tag{142}$$

where N_0 is the number of zeroes of $f(z)$ enclosed by the contour C, and N_p is the number of poles enclosed by C.

Let us now apply this idea to the eigenvalues that are encoded as zeroes in the eqtns (140a) & (140b). Notice that in each of these eqtns, the only dependence on L_z is in the first term in each; we are therefore not interested in the zeroes in the 2nd term (which will give an in general infinite contribution to the sum). To take advantage of these, we can transform the zeroes in

these 2nd term into poles by dividing everything by the 2nd, L_2-independent term in the eqtn in (140), and define the functions *

$$\left.\begin{aligned} f_1(w) &= \frac{[K_1(w)+K_c(w)][K_2(w)+K_c(w)]}{[K_1(w)+K_c(w)][K_2(w)+K_c(w)]}\, e^{2K_c(w)L_2} - 1 \\ f_2(w) &= \frac{[\varepsilon_c(w)K_1(w)+\varepsilon_1(w)K_c(w)][\varepsilon_c(w)K_2(w)+\varepsilon_2(w)K_c(w)]}{[\varepsilon_c(w)K_1(w)-\varepsilon_1(w)K_c(w)][\varepsilon_c(w)K_2(w)-\varepsilon_2(w)K_c(w)]}\, e^{2K_c(w)L_2} - 1 \end{aligned}\right\} \quad (141)$$

which now has zeroes whose position depends implicitly on L_2, and poles whose position does not. Now let's consider the function

$$\left.\begin{aligned} &\oint_C \frac{dw}{2\pi i}\, w \ln F(w,L_2) \qquad \oint_C \frac{dw}{2\pi i}\, w \left[\frac{f_1'(w)}{f_1(w)} + \frac{f_2'(w)}{f_2(w)}\right] \\ &\qquad = \left[\sum_{n_1} w_{n_1}(k_\perp) + \sum_{n_2} w_{n_2}(k_\perp)\right] - \text{pole contributions} \end{aligned}\right\} \quad (142)$$

where, as noted above, the pole contributions are independent of L_2, and we have multiplied by w in the integral to convert a sum over the number of zeroes to a sum over the energies of these zeroes. **

Putting this back into our formula in (141), we now just do the integrations, to get a formula for the Casimir energy given by

$$\bar{E}_{Cas}(L_2) = \frac{\hbar}{8\pi^2}\int_0^\infty k_\perp dk_\perp \oint_C dw \left[w \ln F(w,L_2)\right] + \text{const.} \quad (143)$$

where $F(w) = f_1(w) + f_2(w)$, with these 2 functions defined in (141). If we now differentiate this w.r.t. L_2, we find the Casimir force between the 2 interfaces to be

$$\bar{F}_{Cas}(L_2) = -\frac{\partial \bar{E}_{Cas}(L_2)}{\partial L_2} = -\frac{\hbar}{8\pi^2}\int k_\perp dk_\perp \oint_C dw \left[w \frac{d}{dL_2} F(w,L_2)\right] \quad (144)$$

---

* Here we have $K_\alpha^2(w) = k_\perp^2 - \varepsilon_\alpha(w)w^2/c^2$, as per above, with $\alpha = 1,2,c$.

** Alternatively, we could have defined functions $g_1(w)$, $g_2(w)$ in which we divided the 2nd term by the 1st in (140), and we would then sum over poles, not zeroes