Introduction to Quantum Field Theory

John Cardy

Hilary Term 2012 – Version 19/1/12

Abstract

These notes are intended to supplement the lecture course 'Field Theory in Condensed Matter' and are not intended for wider distribution. Any errors or obvious omissions should be communicated to me at `j.cardy1@physics.ox.ac.uk`.

Contents

1 A Brief History of Quantum Field Theory

Quantum field theory (QFT) is a subject which has evolved considerably over the years and continues to do so. From its beginnings in elementary particle physics it has found applications in many other branches of science, in particular condensed matter physics but also as far afield as biology and economics. In this course we shall be adopting an approach (the *path integral*) which was not the original one, but became popular, even essential, with new advances in the 1970s. However, to set this in its context, it is useful to have some historical perspective on the development of the subject (dates are only rough).

- 19th C. Maxwell's equations – a classical field theory for electromagnetism.
- 1900: Planck hypothesises the photon as the quantum of radiation.

- 1920s/30s: development of particle quantum mechanics: the same rules when applied to the Maxwell field predict photons. However relativistic particle quantum mechanics has problems (negative energy states.)
- 1930s/40s: realisation that relativity + quantum mechanics, in which particles can be created and destroyed, needs a many-particle description where the particles are the quanta of a quantised classical field theory, in analogy with photons.
- 1940s: formulation of the calculation rules for quantum electrodynamics (QED) – Feynman diagrams; the formulation of the path integral approach.
- 1950s: the understanding of how to deal with the divergences of Feynman diagrams through *renormalisation*; QFT methods begin to be applied to other many-body systems *eg* in condensed matter.
- 1960s: QFT languishes – how can it apply to weak + strong interactions?
- 1970s: renormalisation of non-Abelian gauge theories, the renormalisation group (RG) and asymptotic freedom; the formulation of the Standard Model
- 1970s: further development of path integral + RG methods: applications to critical behaviour.
- 1970s: non-perturbative methods, lattice gauge theory.
- 1980s: string theory + quantum gravity, conformal field theory (CFT); the realisation that all quantum field theories are only effective over some range of length and energy scales, and those used in particle physics are no more fundamental than in condensed matter.
- 1990s/2000s: holography and strong coupling results for gauge field theories; many applications of CFT in condensed matter physics.

Where does this course fit in?

In 16 lectures, we cannot go very far, or treat the subject in much depth. In addition this course is aimed at a wide range of students, from experimental particle physicists, through high energy theorists, to condensed matter physicists (with maybe a few theoretical chemists, quantum computing types and mathematicians thrown in). Therefore all I can hope to do is to give you some of the basic ideas, illustrated in their most simple contexts. The hope is to take you all from the Feynman path integral, through a solid grounding in Feynman diagrams, to renormalisation and the RG. From there hopefully you will have enough background to understand Feynman diagrams and their uses in particle physics, and have the basis for understanding gauge theories as well as applications of field theory and RG methods in condensed matter physics.

2 The Feynman path integral in particle quantum mechanics

In this lecture we will recall the Feynman path integral for a system with a single degree of freedom, in preparation for the field theory case of many degrees of freedom.

Consider a non-relativistic particle of unit mass moving in one dimension. The coordinate operator is $\hat{q}$, and the momentum operator is $\hat{p}$. (I'll be careful to distinguish operators and c-numbers.) Of course $[\hat{q}, \hat{p}] = i\hbar$. We denote the eigenstates of $\hat{q}$ by $|q'\rangle$, thus $\hat{q}|q'\rangle = q'|q'\rangle$, and $\langle q'|q''\rangle = \delta(q' - q'')$.

Suppose the hamiltonian has the form $\hat{H} = \frac{1}{2}\hat{p}^2 + V(\hat{q})$ (we can consider more general forms – see later.) The *classical* action corresponding to this is

$$S[q] = \int_{t_i}^{t_f} \left[\tfrac{1}{2}\dot{q}^2 - V(q(t))\right] \, dt$$

where $q(t)$ is a possible classical trajectory, or path. According to Hamilton's principle, the actual classical path is the one which extremises S – this gives Lagrange's equations.

The quantum amplitude for the particle to be at q_f at time t_f given that

it was at q_i at time t_i is

$$M = \langle q_f | e^{-i\hat{H}(t_f - t_i)/\hbar} | q_i \rangle \, .$$

According to Feynman, this amplitude is equivalently given by the path integral

$$I = \int [dq] \, e^{iS[q]/\hbar}$$

which is a integral over all functions (or paths) $q(t)$ which satisfy $q(t_i) = q_i$, $q(t_f) = q_f$. Obviously this needs to be better defined, but we will try to make sense of it as we go along.

In order to understand why this might be true, first split the interval (t_i, t_f) into smaller pieces

$$(t_f, t_{n-1}, \ldots, t_{j+1}, t_j, \ldots, t_1, t_i)$$

with $t_{j+1} - t_j = \Delta t$. Our matrix element can then be written

$$M = \langle q_f | \overbrace{e^{-i\hat{H}\Delta t/\hbar} \ldots e^{-i\hat{H}\Delta t/\hbar}}^{N \text{ factors}} | q_i \rangle$$

(Note that we could equally well have considered a time-dependent hamiltonian, in which case each factor would be different.) Now insert a complete set of eigenstates of $\hat{q}$ between each factor, *eg* at time-slice t_j insert

$$\int_{-\infty}^{\infty} dq(t_j) |q(t_j)\rangle\langle q(t_j)|$$

so that

$$M = \prod_j \int dq(t_j) \langle q(t_{j+1}) | e^{-i\hat{H}\Delta t/\hbar} | q(t_j) \rangle$$

On the other hand, we can think of doing the path integral $\int [dq]$ by first fixing the values $\{q(t_j)\}$ at times $\{t_j\}$ (see Fig. 1) and doing the integrals over the intermediate points on the path, and then doing the integral over the $\{q(t_j)\}$. Thus

$$I = \prod_j \int dq(t_j) \int [dq(t)] \, e^{(i/\hbar) \int_{t_j}^{t_{j+1}} \left(\frac{1}{2}\dot{q}^2 - V(q(t)) \right) dt}$$

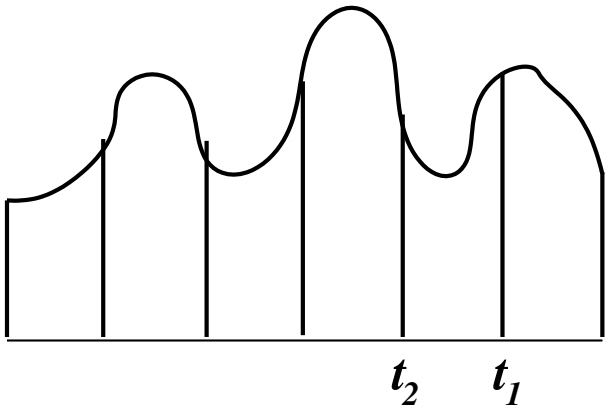

Figure 1: We can imagine doing the path integral by first fixing the values of $q(t)$ at times $(t_1, t_2, \ldots)$.

Thus we can prove that $M = I$ in general if we can show that

$$\langle q(t_{j+1})|e^{-i\hat{H}\Delta t/\hbar}|q(t_j)\rangle = \int [dq(t)]\, e^{(i/\hbar)\int_{t_j}^{t_{j+1}} \left(\frac{1}{2}\dot{q}^2 - V(q(t))\right)dt}$$

for an arbitrarily short time interval Δt. First consider the case when $V = 0$. The path integral is

$$\int [dq] e^{(i/2\hbar)\int_{t_j}^{t_{j+1}} \dot{q}^2 dt}$$

Let $q(t) = q_c(t) + \delta q(t)$ where $q_c(t)$ interpolates linearly between $q(t_j)$ and $q(t_{j+1})$, that is

$$q_c(t) = q(t_j) + (\Delta t)^{-1}(t - t_j)(q(t_{j+1}) - q(t_j))$$

and $\delta q(t_{j+1}) = \delta q(t_j) = 0$. Then

$$\int_{t_j}^{t_{j+1}} \dot{q}^2 dt = (\Delta t)\left(\frac{q(t_{j+1}) - q(t_j)}{\Delta t}\right)^2 + \int (\delta\dot{q})^2 dt$$

and

$$\int [dq] e^{(i/2\hbar)\int_{t_j}^{t_{j+1}} \dot{q}^2 dt} = e^{i\left(q(t_{j+1}) - q(t_j)\right)^2/2\hbar\Delta t} \int [d(\delta q)] e^{(i/2\hbar)\int (\delta\dot{q})^2 dt}$$

The second factor depends on Δt but not $q(t_{j+1})$ or $q(t_j)$, and can be absorbed into the definition, or normalisation, of the functional integral. The first factor we recognise as the spreading of a wave packet initially

localised at $q(t_j)$ over the time interval Δt. This is given by usual quantum mechanics as

$$\langle q(t_{j+1})|e^{-i\hat{p}^2\Delta t/2\hbar}|q(t_j)\rangle$$

(and this can be checked explicitly using the Schrödinger equation.)

Now we argue, for $V \neq 0$, that if Δt is small the spreading of the wave packet is small, and therefore we can approximate $V(q)$ by (say) $V(q(t_j))$. Thus, as $\Delta t \to 0$,

$$\int [dq] e^{(i/\hbar)\int_{t_j}^{t_{j+1}} \left(\frac{1}{2}\dot{q}^2 - V(q(t))\right)dt} \sim \langle q(t_{j+1})|e^{-i(\Delta t/\hbar)\left(\frac{1}{2}\hat{q}^2 + V(\hat{q})\right)}|q(t_j)\rangle$$

Putting all the pieces together, an integrating over the $\{q(t_j)\}$, we obtain the result we want.

As well as being very useful for all sorts of computations, the path integral also provides an intuitive way of thinking about classical mechanics as a limit of quantum mechanics. As $\hbar \to 0$ in the path integral $\int [dq] e^{iS[q]/\hbar}$, the important paths are those corresponding to *stationary phase*, where $\delta S[q]/\delta q = 0$. Other paths giving rapidly oscillating contributions and therefore are suppressed. This is just Hamilton's principle. In the *semi-classical* limit, the important paths will be those close to the classical one. *Periodic* classical orbits will carry a complex phase which will in general average to zero over many orbits. However if the action of a single orbit is $2\pi\hbar\times$ integer, the phase factor is unity and therefore such orbits will dominate the path integral. This is the Bohr-Sommerfeld quantisation condition.

The path integral is not restricted to hamiltonians of the above form, but is more general. An important case is when $\hat{H}(\hat{a}, \hat{a}^\dagger)$ is expressed in terms of annihilation and creation operators $\hat{a}$ and $\hat{a}^\dagger$ satisfying $[\hat{a}, \hat{a}^\dagger] = 1$. In this case, the path integral is obtained by replacing these by complex-valued functions $a(t)$ and $a^*(t)$:

$$\int [da][da^*] e^{(i/\hbar)\int \left(i\hbar a^* \partial_t a - H(a,a^*)\right)dt}$$

This is called a coherent state path integral. Similar versions exist for hamiltonians depending on quantum spins.

2.1 Imaginary time path integrals and statistical mechanics

Sometimes it is useful to consider matrix elements of the form

$$M = \langle q_f | e^{-\hat{H}(\tau_f - \tau_i)/\hbar} | q_i \rangle \,, \tag{1}$$

that is, without the i. An analogous argument to the above shows that this is given by the path integral

$$\int [dq] e^{-S_E[q]/\hbar} \tag{2}$$

where

$$S_E[q] = \int_{\tau_i}^{\tau_f} (\tfrac{1}{2}\dot{q}^2 + V(q(\tau)))d\tau$$

This is called the 'imaginary time' path integral: if we formally let $t = -i\tau$ in the previous result, we get this answer. For reasons that will become apparent in the field theory generalisation, S_E is usually referred to as the euclidean action. Note that the relative sign of the kinetic and potential terms changes between S and S_E.

One application of this idea is to quantum statistical mechanics. The canonical partition function in general is

$$Z = \mathrm{Tr}\, e^{-\beta \hat{H}}$$

where $\beta = 1/k_B T$. For the model under consideration the trace can be written

$$Z = \int dq_i \langle q_i | e^{-\beta \hat{H}} | q_i \rangle$$

where the matrix element is of the form (1) with $\tau_f - \tau_i = \beta\hbar$. Thus Z is also given by the imaginary time path integral (2) over *periodic* paths satisfying $q(\tau_i + \beta\hbar) = q(\tau_i)$.

Another application is to the computation of the ground state energy E_0. If we insert a complete set of eigenstates of $\hat{H}$ into (1) in the limit $\tau_f - \tau_i \equiv T \to \infty$, the leading term has the form $\sim e^{-E_0 T}$. On the other hand, in (2) this is given by paths $q(\tau)$ which minimise $S_E[q]$. Typically they must satisfy $\dot{q}(\tau) \to 0$ as $\tau \to \pm\infty$. In most cases these have $\dot{q} = 0$ throughout, but other cases are more interesting. In particular this leads to an understanding of quantum-mechanical tunnelling.

The imaginary time path integral (2) may also be though of as a partition function in *classical* statistical mechanics. Suppose that we treat τ as a spatial coordinate, and $q(\tau)$ as the transverse displacement of a stretched elastic string tethered at the points τ_i and τ_f. In addition a force, described by an external potential $V(q)$, acts on the string. The euclidean action

$$S_E[q] = \int (\tfrac{1}{2}m(dq/d\tau)^2 + V(q(\tau)))d\tau$$

(where we have restored the particle mass m in the original problem) can now be thought of as the *potential* energy of the string, the first term representing the bending energy where m is the string tension. The partition function of the string in classical statistical mechanics is

$$Z = \int [dq][dp] e^{-\left(\int \frac{1}{2\rho}p^2 d\tau + S_E[q]\right)/k_B T}$$

where p now means the momentum density and $\int \frac{1}{2\rho}p^2 d\tau$ is the kinetic energy, with ρ being the string's mass per unit length. The integral over p just gives a constant, as in a classical gas, so comparing with (2) we see that the imaginary time path integral actually corresponds to a classical partition function at temperature $k_B T = \hbar$. This is the simplest example of one of the most powerful ideas of theoretical physics:

$\Rightarrow$*Quantum mechanics (in imaginary time)* $\equiv$ *classical statistical mechanics in one higher spatial dimension*$\Leftarrow$

3 Path integrals in field theory

A field theory is a system whose degrees of freedom are distributed throughout space. Since the continuous version of this is a little difficult to grasp initially, consider a discrete regular lattice in D-dimensional space whose sites are labelled by $(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \ldots)$. At each site there is a degree of freedom. Instead of $\hat{q}$ and $\hat{p}$ we use $\hat{\phi}$ and $\hat{\pi}$. Thus

$$\begin{aligned} \hat{q} &\rightarrow (\hat{\phi}(\mathbf{x}_1), \hat{\phi}(\mathbf{x}_2), \ldots) \\ \hat{p} &\rightarrow (\hat{\pi}(\mathbf{x}_1), \hat{\pi}(\mathbf{x}_2), \ldots) \end{aligned}$$

satisfying the canonical commutation relations

$$[\hat{\phi}(\mathbf{x}_j), \hat{\pi}(\mathbf{x}_{j'})] = i\hbar\delta_{jj'}$$

The simplest form of the hamiltonian, generalising our single degree of freedom example, is

$$\hat{H} = \sum_j \hat{h}(\hat{\pi}(\mathbf{x}_j), \hat{\phi}(\mathbf{x}_j)) + \tfrac{1}{2}J \sum_{(jj')} (\hat{\phi}(\mathbf{x}_j) - \hat{\phi}(\mathbf{x}_{j'}))^2$$

where the last term couples the degrees of freedom on neighbouring sites. We can take $\hat{h}$ to have the same form as before,

$$\hat{h}(\hat{\pi}(\mathbf{x}_j), \hat{\phi}(\mathbf{x}_j)) = \tfrac{1}{2}\hat{\pi}(\mathbf{x}_j)^2 + V(\hat{\phi}(\mathbf{x}_j))$$

In the path integral version the operators $\hat{\phi}(\mathbf{x}_j)$ are replaced by c-number variables $\phi(\mathbf{x}_j, t)$:

$$\int \prod_j [d\phi(\mathbf{x}_j, t)]\, e^{(i/\hbar)S[\{\phi(\mathbf{x}_j,t)\}]}$$

where

$$S = \int \left(\sum_j (\tfrac{1}{2}\dot{\phi}(\mathbf{x}_j, t)^2 - V(\phi(\mathbf{x}_j, t))) - \tfrac{1}{2}J \sum_{(jj')} (\phi(\mathbf{x}_j, t) - \phi(\mathbf{x}_{j'}, t))^2 \right) dt$$

This is the action for a lattice field theory.

However we are interested in the continuum limit, as the lattice spacing $a \to 0$. The *naive* continuum limit is obtained by replacing sums over lattice sites by integrals:

$$\sum_j \to \int \frac{d^D x}{a^D}$$

and making a gradient (Taylor) expansion of finite differences:

$$\sum_{(jj')} (\phi(\mathbf{x}_j,) - \hat{\phi}(\mathbf{x}_{j'}, t))^2 \to \int \frac{d^D x}{a^D}\, a^2 (\nabla\phi(\mathbf{x}, t))^2$$

After rescaling $\phi \to J^{-1/2} a^{(D-2)/2}\phi$ (and also t), the action becomes

$$S = \int dt d^D x \left(\tfrac{1}{2}\dot{\phi}^2 - \tfrac{1}{2}(\nabla\phi)^2 - V(\phi) \right)$$

This is the action for a classical field theory. The quantum theory is given by the path integral over fields $\phi(\mathbf{x}, t)$

$$\int [d\phi(\mathbf{x}, t)]\, e^{iS[\phi]/\hbar}$$

However, this begs the question of whether this has a meaningful limit as $a \to 0$. The naive answer is no, and making sense of this limit requires the understanding of renormalisation.

3.1 Field theory action functionals

The example that we discussed above has several nice properties:

- it is *local* : this means that S can be written as $\int \mathcal{L}(\phi, \dot{\phi}, \nabla\phi) dt d^D x$ where the lagrangian density depends on the local value of the field and its derivatives. Moreover (more technically) it depends on derivatives only up to second order. It can be shown that higher order derivatives in t lead to violations of causality.
- it is relativistically invariant (with $c = 1$): in 4-vector (or D+1-vector) notation $\mathcal{L}$ can be written

  $$\mathcal{L} = \tfrac{1}{2}(\partial_0\phi)^2 - \tfrac{1}{2}\sum_i (\partial_i\phi)^2 - V(\phi) = \tfrac{1}{2}(\partial_\mu\phi)(\partial^\mu\phi) - V(\phi)$$

  so that if ϕ transforms as a Lorentz scalar, $\mathcal{L}$ is Lorentz invariant. This is of course a requirement for a field theory describing relativistic particles. Another example is

  $$\mathcal{L} = \tfrac{1}{4} F_{\mu\nu} F^{\mu\nu}$$

  where $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$ and A_μ is a Lorentz vector. This is the lagrangian for the electromagnetic field. However in condensed matter physics applications, relativistic invariance is not necessary (although it sometimes emerges anyway, with c replaced by the Fermi velocity or the speed of sound.) Note also that the imaginary time version of the action for our scalar field theory is

  $$S_E = \int \left(\tfrac{1}{2} \sum_{i=1}^{d} (\partial_i\phi)^2 + V(\phi) \right) d^d x$$

where $d = D + 1$ and $d^d x = d^D x d\tau$. That is, τ plays the same role as a spatial coordinate and the theory is invariant under rotations in d-dimensional euclidean space. For this reason the imaginary time versions are called euclidean quantum field theories.

- $\mathcal{L}$ should be invariant under any internal symmetries of the theory. If this is $\phi \to -\phi$, for example, then V should be an even function. In the case of electromagnetism, the symmetry is local gauge invariance.
- the theory be *renormalisable* (see later – although non-renormalisable theories also play a role nowadays.)

3.2 The generating functional

One difference between single particle quantum mechanics and quantum field theory is that we are not usually interested in transition amplitudes between eigenstates $|\phi(\mathbf{x})\rangle$ of the field itself, as the field itself is not physically measurable. In fact, since we usually consider the limit of infinite space, on relativistic grounds we should also consider infinite times. Thus the only meaningful path integral would seem to be

$$\int [d\phi] e^{(i/\hbar)\int_{-\infty}^{\infty} dt \int \mathcal{L} d^D x} \tag{3}$$

which is just a number. In fact, if we consider the euclidean version of this,

$$\int [d\phi] e^{-(1/\hbar)\int_{-\infty}^{\infty} d\tau \int \mathcal{L} d^D x} \tag{4}$$

and relate this to a matrix element between eigenstates $|n\rangle$ of $\hat{H}$, we get

$$\lim_{\tau_f - \tau_i \to \infty} \sum_n e^{-E_n(\tau_f - \tau_i)} \langle n|n\rangle \sim e^{-E_0(\tau_f - \tau_i)} \langle 0|0\rangle$$

Thus we see that (4) (and, by careful definition through analytic continuation, see later, (3)) just tells about the vacuum $\to$ vacuum amplitude, and is thus not very interesting (at least in flat space.)

In order to get any interesting physics we have to 'tickle' the vacuum, by adding sources which can make things happen. The simplest and most

useful way of doing this is to add a source coupling locally to the field itself, that is change the action to

$$S \to S + \int J(x)\phi(x)d^dx$$

The vacuum amplitude is now a functional of this source function $J(x)$:

$$Z[J] = \int [d\phi] e^{iS + i\int J(x)\phi(x)d^dx}$$

We are now using $x = (\mathbf{x}, t)$ to represent a point in Minkowski space (or $(\mathbf{x}, \tau)$ in euclidean space), and we have started using units where $\hbar = 1$, both standard conventions in QFT. It is straightforward to put the right factors back in when we calculate a physical quantity.

Since the i makes this rather ill-defined, we shall, for the time being, develop the theory in the euclidean version

$$Z[J] = \int [d\phi] e^{-S + \int J(x)\phi(x)d^dx}$$

Interesting physical quantities are found by taking functional derivatives of $Z[J]$ with respect to J. For example

$$\frac{1}{Z[0]} \frac{\delta Z[J]}{\delta J(x_1)}\bigg|_{J=0} = \frac{1}{Z[0]} \int [d\phi]\phi(x_1)\, e^{-S[\phi]}$$

By analogy with statistical mechanics in d dimensions, this can be thought of as an expectation value $\langle \phi(x_1) \rangle$. Similarly

$$\frac{1}{Z[0]} \frac{\delta^2 Z[J]}{\delta J(x_1)\delta J(x_2)}\bigg|_{J=0} = \frac{1}{Z[0]} \int [d\phi]\phi(x_1)\phi(x_2)\, e^{-S[\phi]} = \langle \phi(x_1)\phi(x_2) \rangle \,,$$

a correlation function.

But what do these mean in the operator formulation? To see this imagine inserting a complete set of eigenstates. Then as $\tau_i \to \infty$ and $\tau_f \to +\infty$,

$$\int [d\phi]\phi(x_1)\, e^{-S[\phi]} \sim e^{-E_0(\tau_f - \tau_1)}\, e^{-E_0(\tau_1 - \tau_i)}\, \langle 0|\hat{\phi}(x_1)|0\rangle$$

and the first two factors get cancelled by $Z[0]$. Similarly the two-point function is

$$\langle \phi(x_1)\phi(x_2) \rangle = \langle 0|\hat{\phi}(\mathbf{x}_1) e^{-(\hat{H} - E_0)(\tau_1 - \tau_2)} \hat{\phi}(\mathbf{x}_2)|0\rangle$$

where we have emphasised that $\hat{\phi}$, in the Schrödinger picture, depends on the spatial coordinates $\mathbf{x}$ but not τ. However if we go to the Heisenberg picture and define

$$\hat{\phi}(x) = e^{-(\hat{H}-E_0)\tau}\,\hat{\phi}(\mathbf{x})\,e^{(\hat{H}-E_0)\tau}$$

the rhs becomes

$$\langle 0|\hat{\phi}(x_1)\hat{\phi}(x_2)|0\rangle\,.$$

However this is correct only if $\tau_1 > \tau_2$. If the inequality were reversed we would have had to write the factors in the reverse order. Thus we conclude that

$$\frac{1}{Z[0]}\frac{\delta^2 Z[J]}{\delta J(x_1)\delta J(x_2)}\bigg|_{J=0} = \langle\phi(x_1)\phi(x_2)\rangle = \langle 0|\mathbf{T}\left[\hat{\phi}(x_1)\hat{\phi}(x_2)\right]|0\rangle$$

where $\mathbf{T}$ arranges the operators in order of decreasing τ.

$\Rightarrow$*Functional derivatives of* $Z[J]$ *give vacuum expectation values of time-ordered products of field operators*$\Leftarrow$

This result continues to hold when we go back to real time t. Fortunately it is precisely these vacuum expectation values of time-ordered products which arise when we do scattering theory.

In field theory, the correlation functions are also called Green functions (as we'll see, for a free field theory they are Green functions of differential operators), or simply the N-point functions

$$G^{(N)}(x_1,\ldots,x_N) = \langle\phi(x_1)\ldots\phi(x_N)\rangle = \frac{1}{Z[0]}\,\frac{\delta^N Z[J]}{\delta J(x_1)\ldots\delta J(x_N)}\bigg|_{J=0}$$

Equivalently

$$\frac{Z[J]}{Z[0]} = \sum_{N=0}^{\infty}\frac{1}{N!}\int d^dx_1\ldots\int d^dx_N\,G^{(N)}(x_1,\ldots,x_N)J(x_1)\ldots J(x_N)$$

$Z[J]$ is called the *generating function* for the N-point functions.

It is also useful to define

$$W[J] \equiv \log Z[J]\,,$$

which is analogous to the free energy in statistical mechanics. We expect, by analogy, that $W[0]$ is proportional to the total space-time volume VT, and that, if the sources J are localised to a finite region of space-time, that $W[J] - W[0]$ is finite in the limit $VT \to \infty$. Thus functional derivatives of W wrt J should also be finite. These give what are called the *connected* correlation functions $\langle \phi(x_1) \dots \phi(x_N) \rangle_c$ or $G^{(N)}(x_1, \dots, x_N)_c$. The reason for this will become apparent when we write them in terms of Feynman diagrams. For example

$$\left. \frac{\delta^2 W[J]}{\delta J(x_1) \delta J(x_2)} \right|_{J=0} = \langle \phi(x_1)\phi(x_2) \rangle_c = \langle \phi(x_1)\phi(x_2) \rangle - \langle \phi(x_1) \rangle \langle \phi(x_2) \rangle$$

$W[J]$ is the generating function for the connected N-point functions.

3.3 The propagator in free field theory

The only path integrals we can actually do (except in certain esoteric theories with supersymmetry) are gaussian, that is when the action S is at most quadratic in the field ϕ. However this is an important case, corresponding to a free field theory. As usual, we consider the euclidean case first.

$$Z_0[J] = \int [d\phi]\, e^{-\int \left[\frac{1}{2}(\partial\phi)^2 + \frac{1}{2}m^2\phi^2\right] d^dx + \int J(x)\phi(x) d^dx}$$

(So far m is just a parameter, but it will turn out that in Minkowski space this theory describes free relativistic particles of mass m.) The first term can be integrated by parts to give $\frac{1}{2}\phi(-\partial^2)\phi$.

Define Fourier transforms:

$$\begin{aligned} \tilde{\phi}(p) &= \int d^dx\, e^{-ip\cdot x}\, \phi(x) \\ \phi(x) &= \int \frac{d^dp}{(2\pi)^d}\, e^{ip\cdot x}\, \tilde{\phi}(p) \end{aligned}$$

and similarly for $\tilde{J}(p)$ and $J(x)$. (Note that in field theory it is conventional to put the factors of 2π as above.)

The negative of the expression in the exponential is then

$$\int \frac{d^dp}{(2\pi)^d} \left[\tfrac{1}{2}\tilde{\phi}(p)(p^2 + m^2)\tilde{\phi}(-p) - \tilde{J}(p)\tilde{\phi}(-p)\right]$$

Completing the square on the expression in square brackets:

$$\frac{1}{2}\left[\tilde{\phi}(p)-\frac{1}{p^2+m^2}\tilde{J}(p)\right](p^2+m^2)\left[\tilde{\phi}(-p)-\frac{1}{p^2+m^2}\tilde{J}(-p)\right]$$
$$-\frac{1}{2}\tilde{J}(p)\frac{1}{p^2+m^2}\tilde{J}(-p)$$

Now the functional integral $\int[d\phi(x)]$ can equally well be carried out over $\int[d\tilde{\phi}(p)]$. Shifting the integration variable $\tilde{\phi}(p)=\tilde{\phi}(p)'+(p^2+m^2)^{-1}\tilde{J}(p)$ gives

$$Z_0[J]=\int[d\tilde{\phi}']\,e^{-\frac{1}{2}\int(d^dp/(2\pi)^d)\tilde{\phi}(p)'(p^2+m^2)\tilde{\phi}(-p)'+\frac{1}{2}\int(d^dp/(2\pi)^d)\tilde{J}(p)(p^2+m^2)^{-1}\tilde{J}(-p)}$$

The first term in the exponential gives a factor independent of J, so

$$Z_0[J]=Z_0[0]\,e^{\frac{1}{2}\int(d^dp/(2\pi)^d)\tilde{J}(p)(p^2+m^2)^{-1}\tilde{J}(-p)}$$

Going back to coordinate space

$$Z_0[J]=Z_0[0]\,e^{\frac{1}{2}\int d^dx'\int d^dx''J(x')\Delta(x'-x'')J(x'')} \tag{5}$$

where

$$\Delta(x'-x'')\equiv\int\frac{d^dp}{(2\pi)^d}\frac{e^{ip\cdot(x'-x'')}}{p^2+m^2}\,.$$

With this result in hand we can now compute correlation functions in the free theory, *e.g.*

$$\langle\phi(x_1)\rangle_0=\tfrac{1}{2}\int d^dx''\Delta(x_1-x'')J(x'')+\tfrac{1}{2}\int d^dx'\Delta(x'-x_1)J(x')\Big|_{J=0}=0$$

$$\langle\phi(x_1)\phi(x_2)\rangle_0=\tfrac{1}{2}\Delta(x_1-x_2)+\tfrac{1}{2}\Delta(x_2-x_1)=\Delta(x_1-x_2)$$

$\Delta(x_1-x_2)$ is thus the 2-point function $G_0^{(2)}(x_1,x_2)$ in the free theory.

$\langle\phi(x_1)\dots\phi(x_N)\rangle_0=0$ if N is odd, in this theory, because of a symmetry of the lagrangian under $\phi(x)\to-\phi(x)$, but, for example

$$\begin{aligned}\langle\phi(x_1)\phi(x_2)\phi(x_3)\phi(x_4)\rangle_0 \;=\;&\Delta(x_1-x_2)\Delta(x_3-x_4)\\&+\Delta(x_1-x_3)\Delta(x_2-x_4)\\&+\Delta(x_1-x_4)\Delta(x_2-x_3)\,.\end{aligned}$$

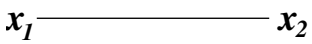

Figure 2: Graphical representation of the propagator $\Delta(x_1 - x_2)$.

Figure 3: Wick contractions for the 4-point function. Each line represents a factor Δ.

To see this, imagine expanding the exponential in (5) to $O(J^4)$. We get a non-zero contribution if each of (x_1, x_2, x_3, x_4) hits one of the integration variables x' or x''. In general, for N even,

$$\langle \phi(x_1) \dots \phi(x_N) \rangle_0 = \sum \Delta(x_{j_1} - x_{j'_1}) \dots \Delta(x_{j_{N/2}} - x_{j'_{N/2}}) \,,$$

where the sum is over all distinct ways of grouping the set $\{1, 2, \dots, N\}$ into pairs. This result, which in fact holds for any gaussian integral, is the path integral version of *Wick's theorem*. It tells us that in the free theory, every correlation function can be expressed in terms of $G_0^{(2)}$. Another way of stating it is to observe that the generating function for connected correlation functions $W[J]$ is quadratic in J. Thus all connected N-point functions vanish for $N > 2$.

At this stage we can begin to introduce a graphical notation which will become one of the building blocks for Feynman diagrams. We denote $\Delta(x_1 - x_2)$ by an (unoriented) line connecting the points x_1 and x_2, as shown in Fig. 2. (it doesn't matter exactly where we put the points, only the topology is important.) Then Wick's theorem for $N = 4$ can be expressed by connecting up the points (x_1, x_2, x_3, x_4) by lines in all possible ways, such that exactly one line ends at each point. See Fig.3.

3.3.1 Minkowski space

In real time, the path integral is less well-defined, because the integrand is oscillating rather than exponentially damped at large values of ϕ:

$$Z_0[J] = \int [d\phi]\, e^{i \int \left(\frac{1}{2}(\partial_\mu \phi)(\partial^\mu \phi) - \frac{1}{2} m^2 \phi^2 \right) dt d^D \mathbf{x} + i \int J \phi dt d^D \mathbf{x}}$$

where $(\partial_\mu\phi)(\partial^\mu\phi) = (\partial_t\phi)^2 - (\nabla\phi)^2$.

One way to make this better defined is to give the parameter m^2 a small negative imaginary part

$$m^2 \to m^2 - i\epsilon$$

Now that the integral is absolutely convergent we can rotate the contour in the t-integration by letting $t = -i\tau$, whereby

$$Z_0[J] = \int [d\phi]\, e^{-\int \left(\frac{1}{2}((\partial_\tau\phi)^2 + (\nabla\phi)^2) + \frac{1}{2}(m^2 - i\epsilon)\phi^2\right) d^dx + \int J\phi d^dx}$$

This is the generating function in euclidean space. So we can get all the results in Minkowski space by substituting $\tau = it$ in their euclidean versions. This technique is called *Wick rotation*. Note that when we do this,

$$p \cdot x = p_0\tau + \mathbf{p}\cdot\mathbf{x} \to ip_0 t + \mathbf{p}\cdot\mathbf{x}$$

so that we have to let $p_0 \to ip_0$ and then

$$p \cdot x \,(\text{euclidean}) \to -p_\mu x^\mu \,(\text{Minkowski})$$

Thus

$$\begin{aligned} Z_0[J] &= Z_0[0]\, e^{-(i/2)\int (d^dp/(2\pi)^d)\widetilde{J}(p)(p^2 - m^2 + i\epsilon)^{-1}\widetilde{J}(-p)} \\ &= e^{-(i/2)\int d^dx' d^dx'' J(x')\Delta_F(x' - x'')J(x'')} \end{aligned}$$

where

$$\Delta_F(x_1 - x_2) = \int \frac{d^dp}{(2\pi)^d} \frac{e^{-ip_\mu(x_1^\mu - x_2^\mu)}}{p^2 - m^2 + i\epsilon} \tag{6}$$

Δ_F is called the *Feynman propagator*. We shall discuss its physical interpretation below. If we recall that

$$G^{(2)}(x_1, x_2) = \frac{\delta^2 Z[J]}{\delta(iJ(x_1))\delta(iJ(x_2))}$$

(note the factors of i), we see that

$$G_0^{(2)}(x_1, x_2) = i\Delta_F(x_1 - x_2)$$

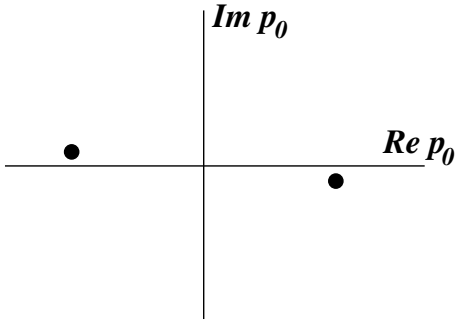

Figure 4: Poles of the Feynman propagator in the complex p_0-plane.

Let us examine the p_0 integration in (6):

$$\int_{-\infty}^{\infty} \frac{e^{-ip_0(t_1-t_2)}}{p_0^2 - \mathbf{p}^2 - m^2 + i\epsilon} \frac{dp_0}{2\pi}$$

The integrand has poles at $p_0 = \pm\sqrt{\mathbf{p}^2 + m^2 - i\epsilon}$ (see Fig. 4). Suppose that $t_1 > t_2$. Then we can close the p_0 contour in the lower half plane, picking up the pole with the positive sign of Re p_0. This gives

$$G_0^{(2)}(x_1, x_2) = \int \frac{d^D p}{(2\pi)^D} \frac{e^{-i\sqrt{\mathbf{p}^2+m^2}(t_1-t_2)+i\mathbf{p}\cdot(\mathbf{x}_1-\mathbf{x}_2)}}{2\sqrt{\mathbf{p}^2+m^2}}$$

If $t_2 > t_1$ we pick up the other pole and get the same result with t_1 and t_2 interchanged. Now recall that in the operator formulation is a vacuum expectation value of a time-ordered product:

$$G_0^{(2)}(x_1, x_2) = \langle 0|\mathbf{T}[\hat{\phi}(x_1)\hat{\phi}(x_2)]|0\rangle$$

Thus if we define

$$\tilde{\hat{\phi}}(\mathbf{p}, t) = \int d^D x\, e^{-i\mathbf{p}\cdot\mathbf{x}} \hat{\phi}(\mathbf{x}, t)$$

we see that, for $t_1 > t_2$,

$$\langle 0|\tilde{\hat{\phi}}(\mathbf{p_1}, t_1)\tilde{\hat{\phi}}(\mathbf{p_2}, t_2)^\dagger|0\rangle = (2\pi)^D \delta^{(D)}(\mathbf{p}_1 - \mathbf{p}_2) \frac{e^{-i\sqrt{\mathbf{p}_2^2+m^2}(t_1-t_2)}}{2\sqrt{\mathbf{p}_2^2+m^2}}$$

The interpretation of this is that $\tilde{\hat{\phi}}(\mathbf{p_2}, t_2)^\dagger$ creates a particle of momentum $\mathbf{p}_2$ and energy $\sqrt{\mathbf{p}_2^2 + m^2}$ at time t_2, and $\tilde{\hat{\phi}}(\mathbf{p_1}, t_1)$ destroys a particle (note that since ϕ is real, $\tilde{\hat{\phi}}(\mathbf{p_2}, t_2)^\dagger = \tilde{\hat{\phi}}(-\mathbf{p_2}, t_2)$. The rhs is the quantum amplitude for the particle to propagate from x_2 to x_1, and is therefore called the *propagator*.

4 Interacting field theories

Interactions in field theories, leading to nontrivial particle scattering, correspond to the appearance in the lagrangian density of terms higher than quadratic in the fields. In general such QFTs are not exactly solvable, and it is necessary to adopt various approximation schemes. The most common is perturbation theory in the higher order terms. This leads to Feynman diagrams.

4.1 Feynman diagrams

In general we write the lagrangian density as

$$\mathcal{L} = \mathcal{L}_0 + \mathcal{L}_I$$

where $\mathcal{L}_I$ contains only terms quadratic in the fields. We first consider the euclidean case. Consider the N-point function

$$G^{(N)}(y_1, \ldots, y_N) = \frac{\int [d\phi] \phi(y_1) \ldots \phi(y_N) e^{-S_0 - \int \mathcal{L}_I(x) d^d x}}{\int [d\phi] e^{-S_0 - \int \mathcal{L}_I(x) d^d x}} \tag{7}$$

where $\mathcal{L}_I(x) \equiv \mathcal{L}_I[\phi(x)]$. We can expand the numerator and denominator as a power series in $\mathcal{L}_I$ by writing

$$e^{-\int \mathcal{L}_I(x) d^d x} = \sum_{n=0}^{\infty} \frac{(-1)^n}{n!} \int \mathcal{L}_I(x_1) \ldots \mathcal{L}_I(x_N) d^d x_1 \ldots d^d x_n$$

Each term in the numerator then involves evaluating

$$\langle \phi(y_1) \ldots \phi(y_N) \, \mathcal{L}_I(x_1) \ldots \mathcal{L}_I(x_n) \rangle_0 \tag{8}$$

in the free theory, and, since $\mathcal{L}_I$ is a polynomial in the field ϕ, this can be done using Wick's theorem.

Throughout this course, with a few exceptions, we shall consider in detail one of the simplest interacting field theories, with

$$\mathcal{L}_I = \frac{\lambda}{4!} \phi(x)^4 \tag{9}$$

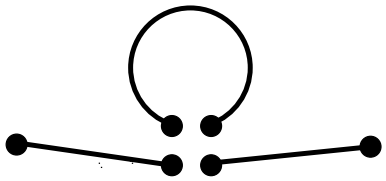

Figure 5: One of the possible Wick contractions for the connected first order term in 2-point function.

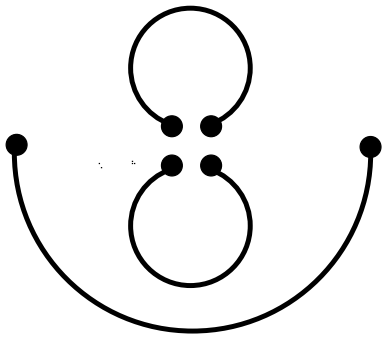

Figure 6: One of the possible Wick contractions for the disconnected first order term in 2-point function.

Although this theory has few direct applications in particle physics (except maybe to describe the self-interactions of Higgs bosons) it does illustrate many of the important aspects of QFT without many of the obscuring algebraic details which arise in theories of fermions and gauge bosons. In addition, it has direct application to describing critical behaviour in magnets and other condensed matter systems.

In order to apply Wick's theorem it is useful to consider the coordinates of the four powers of $\phi(x)$ occurring in (9) as being slightly different. In (8) we then have a collection of fields $\phi(y_j)$ denoted by single external points, and $\phi(x_j)^4$ by quadruplets of internal points. All of these must be connected in pairs by propagators in all possible ways, according to Wick's theorem. It is useful to consider the first few simplest cases.

Consider first the 2-point function $G^{(2)}(y_1, y_2)$. At order $n = 0$ it is given by a single propagator, $\Delta(y_1 - y_2)$. For $n = 1$ there are two different classes of Wick contractions: either each point y_1, y_2 is contracted to one of the points at x, and the remaining two at x are contracted to each other, as in Fig. 5 – this can happen in $4 \cdot 3 = 12$ ways; or y_1 and y_2 are connected to

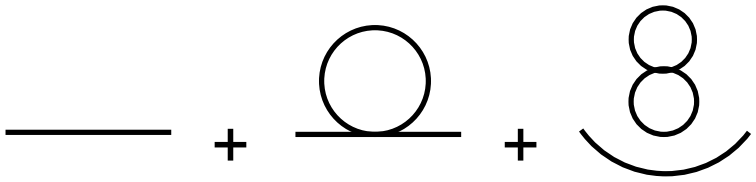

Figure 7: Feynman diagrams for the numerator in the 2-point function to $O(\lambda)$.

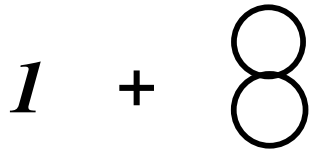

Figure 8: Feynman diagrams for the denominator in the 2-point function to $O(\lambda)$.

each other and the remaining 4 points at x are contracted together, as in Fig. 6 – this happens in 3 ways. Thus, to $O(\lambda)$, the numerator in (7) is

$$\begin{aligned}\Delta(y_1-y_2) \quad & -12\frac{\lambda}{4!}\int \Delta(y_1-x)\Delta(x-x)\Delta(x-y_2)d^dx \\ & -3\frac{\lambda}{4!}\Delta(y_1-y_2)\int \Delta(x-x)^2 d^dx + \cdots\end{aligned}$$

This is represented by the Feynman diagrams in Fig. 7. Each diagram corresponds to a topologically different set of contractions. The last integral appears to diverge proportional to total space-time volume VT, since its integrand in independent of x. However the denominator in (7) is, to the same order, given by the diagrams in Fig. 8

$$1 - 3\frac{\lambda}{4!}\int \Delta(x-x)^2 d^dx + \cdots$$

If we first keep VT finite, divide these two expressions and expand in λ to the order required, the potentially divergent terms cancel. This is an example of a general result that diagrams which contain pieces not connected to the external points can be ignored. Note that the overall numerical factor in the connected diagram is 1/2. This is called its symmetry factor.

The diagrams contributing to $G^{(2)}$ at second order, $n = 2$, are shown in Fig. 9. As an example consider the last one. We can either connect y_1 to x_1 and y_2 to x_2, or vice versa. These give the same result on integration over x_1, x_2. Each of these contractions can be done in $4 \cdot 4 = 16$ ways. see

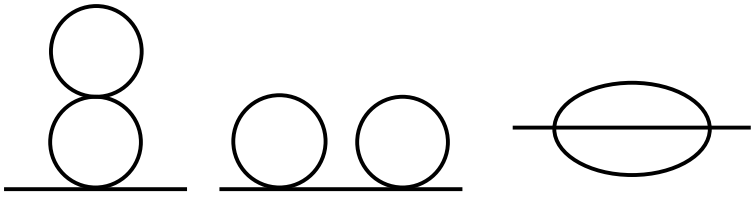

Figure 9: Connected 2-loop diagrams for the 2-point function.

Figure 10: Starting to make the Wick contractions for the last diagram in Fig. 9.

Fig. 10. The remaining 3×3 points can be connected in $(3!)^2/3! = 6$ ways. Thus this diagram evaluates to

$$(2 \cdot 16 \cdot 6)\frac{1}{2!}\left(\frac{-\lambda}{4!}\right)^2 \int \Delta(y_1 - x_1)\Delta(x_1 - x_2)^3\Delta(x_2 - y_2)d^dx_1 d^dx_2$$

In this case the overall factor is $1/3!$.

Now consider the 4-point function. At order $n = 0$ there are only the disconnected diagrams shown in Fig. 3. Since we can build disconnected diagrams out of lower-order connected ones, we need consider only the connected ones. To first order, this is shown in Fig. 11. There are 4! ways to connect the external point to the internal one, which exactly cancels the $1/4!$ in (9) (which is why it was put there in the first place.) Thus the contribution to $G_c^{(4)}$ is

$$-\lambda \int \prod_{j=1}^{4} \Delta(y_j - x)d^dx$$

To order $n = 2$, the connected diagrams are shown in Fig. 12 and Fig. 13.

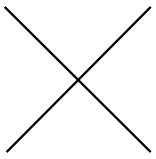

Figure 11: Lowest order diagram for the connected 4-point function.

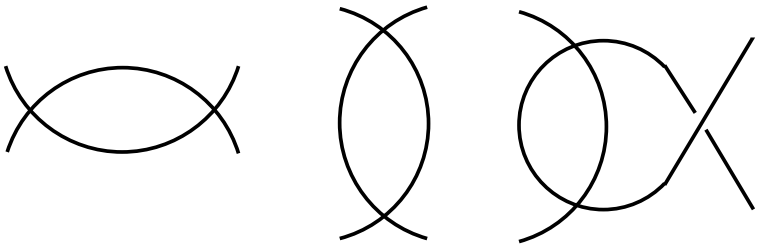

Figure 12: Connected 1-loop 1PR diagrams for the 4-point function.

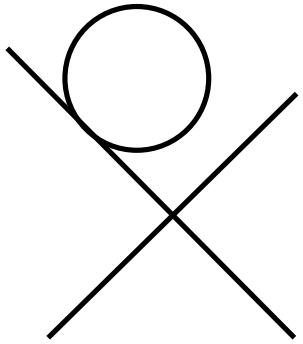

Figure 13: A 1-particle reducible 1-loop diagram for the 4-point function.

The first one in Fig. 12 evaluates to (check the overall factor for yourselves)

$$\frac{(-\lambda)^2}{2}\int \Delta(y_1-x_1)\Delta(y_2-x_1)\Delta(x_1-x_2)^2\Delta(x_2-y_3)\Delta(x_2-y_4)d^dx_1d^dx_2$$

Diagrams like that in Fig. 13 are formed by attaching together lower order diagrams by a single external leg. They are called one-particle reducible (1PR) and we shall see that they can be taken into account automatically.

At this point we are ready to state the

4.1.1 Feynman rules in euclidean position space for the connected N-point functions of ϕ^4 theory

1. draw all topologically distinct connected diagrams with N external lines and each internal vertex attached to 4 lines
2. to each line associate a factor $\Delta(x'-x'')$ (where x' and x'' can be either an internal or external vertex)
3. to each internal vertex associate a factor $-\lambda$

4. integrate over internal vertices $\prod_j \int d^d x_j$

5. multiply by the symmetry factor 1/(integer).

Only the last rule can sometimes cause confusion. For a generic diagram, the number of ways the vertices can be connected up exactly cancels the factor $1/n!$ coming from the expansion of the exponential and the factors $(1/4!)^n$. However in diagrams with some degree of symmetry the number of possible contractions is fewer. This is what happened in our examples above. In the second diagram of Fig. 7 we could imagine twisting the bubble by 180°, and the same in Fig. 12. These lead to a factor 1/2!. Similarly for the first two diagrams of Fig. 9 we get $(1/2!)^2$. In the third diagram there is a permutation symmetry of the 3 external lines, leading to a factor 1/3!. In general the symmetry factor is the inverse of the number of elements in the symmetry group of the diagram. When in doubt, however, you can always go back to enumerating the number of distinct Wick contractions which lead to a given diagram.

4.1.2 Feynman rules in euclidean momentum space

Since the free propagator $\Delta(x' - x'')$ is expressed as an integral over p, and in any case for most purposes we are more interested in the Fourier transforms of N-point functions, it is easier to evaluate most diagrams in momentum space. To do this, we first insert

$$\Delta(x_j - x_k) = \int \frac{d^d p_{jk}}{(2\pi)^d} \frac{e^{ip_{jk}(x_j - x_k)}}{p_{jk}^2 + m^2}$$

for each line. At same time we define the Fourier transforms

$$\begin{aligned} \int G^{(N)}(y_1, \ldots, y_N)_c \, e^{i(p_1 y_1 + \cdots p_N y_N)} d^d y_1 \ldots d^d y_N \\ \equiv \widetilde{G}^{(N)}(p_1, \ldots, p_N)_c \, (2\pi)^d \delta^{(d)}(p_1 + \cdots + p_N) \,, \end{aligned}$$

pulling out an overall momentum-conserving delta-function that must be there since $G^{(N)}$ is a function of only coordinate differences. Then we can carry out the integrations over the internal coordinates. At the jth vertex

we have

$$\int e^{i(\sum_k p_{kj})x_j} d^d x_j = (2\pi)^d \delta^{(d)}(\sum_k p_{kj}) \,,$$

that is, the total momentum entering a given internal vertex vanishes – the momentum is *conserved*. This means that we get linear relations between the different p_{jk} which can be solved in terms of the external momenta p_j. If the diagram contains loops, however, there will remain an number of undetermined integration variable equal to the number of loops This is the smallest number of lines which must be broken to get a connected tree diagram with no loops. We then have the Feynman rules in momentum space for the $\widetilde{G}^{(N)}(p_1, \ldots, p_N)$:

1. draw all topologically distinct connected diagrams with N external lines and each internal vertex attached to 4 lines
2. assign momenta flowing along each line so that the external lines have momenta $\{p_j\}$ and momentum is conserved at each internal vertex
3. to each line associate a factor $(p_{jk}^2 + m^2)^{-1}$
4. to each internal vertex associate a factor $-\lambda$
5. integrate over remaining loop momenta $\prod_j \int d^d p/(2\pi)^d$
6. multiply by the symmetry factor 1/(integer).

In Minkowski space, the combinatorics are the same: only the factors of i differ:

1. draw all topologically distinct connected diagrams with N external lines and each internal vertex attached to 4 lines
2. assign momenta flowing along each line so that the external lines have momenta $\{p_j\}$ and momentum is conserved at each internal vertex
3. to each line associate a factor $i/(p_{jk}^2 - m^2 + i\epsilon)$
4. to each internal vertex associate a factor $i\lambda$
5. integrate over remaining loop momenta $\prod_j \int d^d p/(2\pi)^d$
6. multiply by the symmetry factor 1/(integer).

4.1.3 Feynman rules for other QFTs

The rules for other field theories have basically the same building blocks of propagators and vertices. However these can vary depending on the symmetries of the lagrangian and the form of the interactions. It is important to be able to write down the Feynman rules without always going back to first principles. However this comes only with experience!

As an example, consider a *complex* scalar field $\phi(x)$, with (euclidean) lagrangian density

$$\mathcal{L} = (\partial\phi^*) \cdot (\partial\phi) + m^2\phi^*\phi + \tfrac{1}{4}\lambda(\phi^*\phi)^2$$

Note how the factors of 2 are chosen to differ from the real case. This is because when we work out the 2-point function in the free theory we find

$$\langle\phi(x_1)\phi^*(x_2)\rangle = \int \frac{d^dp}{(2\pi)^d}\frac{e^{ip(x_1-x_2)}}{p^2+m^2} = \Delta(x_1-x_2)$$

as before. On the other hand $\langle\phi\phi\rangle = \langle\phi^*\phi^*\rangle = 0$. This can be traced to a U(1) *symmetry* of the lagrangian under $\phi \to e^{i\alpha}\phi$, $\phi^* \to e^{-i\alpha}\phi^*$. In field theory, such continuous symmetries are associated with conserved currents, in this case

$$J_\mu \propto i(\phi^*\partial_\mu\phi - \phi\partial_\mu\phi^*)$$

You can check, using Lagrange's equations, that $\partial^\mu J_\mu = 0$, and, moreover, using Wick's theorem, that

$$\langle\phi(x_1)\int_S J_\mu dS^\mu\phi(x_2)\rangle \propto \langle\phi(x_1)\phi^*(x_2)\rangle$$

where $\int_S$ is over the surface of a small sphere enclosing x_2 (but not x_1). If J_μ is normalised correctly the coefficient is unity. This means that $\phi^*(x)$ acts as a source of unit flux of the current (and $\phi(x)$ as a sink.) In Minkowski space, the particles created and destroyed by the quantum field $\hat{\phi}(x)$ have unit charge with respect to this U(1) symmetry. Physically, this could be ordinary electric charge or something more exotic. Note that if we interpret $\langle\phi(x_1)\phi^*(x_2)\rangle = \langle\phi(x_2)\phi^*(x_1)\rangle$ in terms of the vacuum expectation value of time-ordered field operators as before, we see that in this case $\hat{\phi}^\dagger$ creates particles of charge $+1$ but it also destroys particles of charge -1 (and

Figure 14: Propagator for a complex scalar field is oriented and indicates flow of charge.

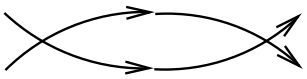

Figure 15: This diagram has symmetry factor 1/2.

oppositely for $\hat{\phi}$). That is, $\langle\phi(x_1)\phi^*(x_2)\rangle$ describes the propagation of a particle of charge $+1$ if $t_1 > t_2$, but also the propagation of a particle of charge -1, the *antiparticle*, if $t_2 > t_1$. This is an example of how, in a relativistic field theory, every particle has to have a corresponding antiparticle of opposite charge.

In Feynman diagrams, we have to distinguish the ends of the propagator according to whether they correspond to ϕ^* or ϕ. We do this by orienting the line with an arrow from x_2 to x_1 (see Fig. 14). The above discussion shows that we can think of the arrow as indicating the flow of charge along the line. An interaction vertex $(\phi^*\phi)^2$ then always has exactly two arrows entering and two leaving: it conserves charge. If we work out the 4-point function $\langle\phi(y_1)\phi(y_2)\phi^*(y_3)\phi^*(y_4)\rangle$ to $O(\lambda)$ we find that the number of allowed contractions exactly cancels the factor $\frac{1}{4}$ in the lagrangian.

Note also that the symmetry factors can change. In Fig. 15 the symmetry factor is $\frac{1}{2}$, but in Fig. 16 it is 1.

4.2 Evaluation of Feynman diagrams

The evaluation of the integrals involved in Feynman diagrams can be quite difficult - in general only one and (some) 2-loop diagrams can be evaluated analytically. However it is important to understand how to do this in order to understand the properties of the result.

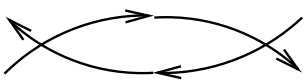

Figure 16: This diagram has symmetry factor 1.

As a first example consider the 1-loop integral for the 2-point function (in euclidean space) corresponding to the second diagram in Fig. 7:

$$I_2 = \int \frac{d^d p}{(2\pi)^d} \frac{1}{p^2 + m^2}$$

(we keep d arbitrary because it will turn out the the dependence on this is interesting.) There are many ways to evaluate this. One is to write

$$(p^2 + m^2)^{-1} = \int_0^\infty e^{-u(p^2+m^2)} du$$

then to perform the p-integration

$$\int e^{-u\sum_{i=1}^d p_i^2} \prod_{i=1}^d dp_i = \left(\frac{\pi}{u}\right)^{d/2}$$

Thus

$$I_2 = \frac{\pi^{d/2}}{(2\pi)^d} \int_0^\infty u^{-d/2} e^{-um^2} du$$

Rescaling $u \to um^{-1/2}$ then gives

$$I_2 = \frac{\pi^{d/2}}{(2\pi)^d} \Gamma(1 - \tfrac{d}{2})\, m^{d-2} \tag{10}$$

This simple example illustrates an important point. The result makes sense for non-integer values of d, even though of course in physical applications it is always an integer. However, the gamma function is infinite at its poles, the first (as d increases) being at $d = 2$. This infinity reflects the fact that for large p the original integral behaves like $\int d^d p/p^2$ and therefore converges only for $d < 2$. Only in this case is the result of the integral given by (10), otherwise the integral diverges. This is an example of the ultraviolet (UV) divergences which plague the perturbation expansion of QFT. The integral could be rendered finite by restricting the allowed values of p to satisfy $|p| < \Lambda$, where Λ is called a UV cut-off. This is an example of a UV *regulator*: something which makes the integrals finite. In the context of the original lattice field theory this would make sense, with $\Lambda \sim a^{-1}$, the inverse lattice spacing. In the critical dimension, in this case two, the leading dependence on Λ is logarithmic. It is easy to see from the original integral

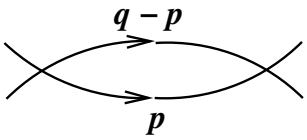

Figure 17: Internal momenta assignment for the 1-loop contribution to the 4-point function. Note that the arrows now indicate flow of momentum, not charge!

that in this case $I \sim (1/2\pi)\log(\Lambda/m)$. Another way to regulate the theory would be to assume that $d < 2$ and then try to continue to the physical dimensionality at the end: this is called dimensional regularisation.

A second more complicated example is the 1-loop contribution to the 4-point function. With the momentum assigned as in Fig. 17 the integral is

$$I_4 = \int \frac{d^d p}{(2\pi)^d} \frac{1}{(p^2+m^2)((q-p)^2+m^2)}$$

where $q = p_1 + p_2 = -p_3 - p_4$.

A useful tool is Feynman's identity

$$\frac{1}{a_1 \dots a_n} = \frac{1}{(n-1)!} \int_{\{x_j \geq 0\}} \frac{\prod_{j=1}^n dx_j \delta(\sum_{j=1}^n x_j - 1)}{[x_1 a_1 + \dots + x_n a_n]^n}$$

where the integral is over an $(n-1)$-dimensional simplex. Applying this for $n = 2$ to our integral we get

$$I = \int_0^1 dx \int \frac{d^d p}{(2\pi)^d} \frac{1}{[x(p^2+m^2) + (1-x)((q-p)^2+m^2)]^2}$$

Expanding out the expression in square brackets and completing the square:

$$p^2 - 2(1-x)p\cdot q + (1-x)q^2 + m^2 =$$
$$(p-(1-x)q)^2 - (1-x)^2 q^2 + (1-x)q^2 + m^2 = p'^2 + x(1-x)q^2 + m^2$$

where $p' = p - (1-x)q$. Then

$$I_4 = \int_0^1 dx \int \frac{d^d p'}{(2\pi)^d} \frac{1}{[p'^2 + x(1-x)q^2 + m^2]^2}$$

The p'-integral can be done using the same method as above:

$$I_4 = \int_0^1 dx \int_0^\infty u du e^{-u(x(1-x)q^2+m^2)} \int \frac{d^d p'}{(2\pi)^d} e^{-up'^2}$$

$$= \frac{\pi^{d/2}}{(2\pi)^d} \int_0^\infty u^{1-d/2} e^{-u} du \int_0^1 (x(1-x)q^2 + m^2)^{d/2-2} dx$$
$$= \frac{\pi^{d/2}}{(2\pi)^d} \Gamma(2 - \tfrac{d}{2}) \int_0^1 (x(1-x)q^2 + m^2)^{d/2-2} dx$$

The integral is of a finite integrand over a finite interval and could easily be done numerically. Note that the first pole is now at $d = 4$, reflecting the fact that the integral behaves like $\int (d^d p/(p^2)^2)$ for large p.

5 Renormalisation

We have seen that most Feynman integrals are UV divergent for large enough d, and therefore do not immediately make sense. However, this perturbation expansion is in powers of a quantity λ which, we shall argue, is not itself directly measurable. Therefore there is no physical requirement for the coefficients in the expansion to be well-defined. The renormalisation procedure attempts to make sense from this nonsense. It proceeds in several steps:

1. first relabel the fields $\phi \to \phi_0$ and the parameters $m \to m_0$, $\lambda \to \lambda_0$ in recognition of the fact that these are not the physical quantities (they are called the *bare* field, mass, and coupling.) Similarly relabel the $G^{(N)}$ as $G_0^{(N)}$ - the bare N-point functions. [Not to be confused with the earlier subscript 0 for the free theory.]

2. understand exactly where the divergences occur.

3. regularise the theory, that is make all Feynman integrals finite. This can be done *e.g.*, by cutting off all internal momentum integrals $|p| < \Lambda$, or by reducing d until all integrals are finite (or otherwise).

4. decide what quantities are physically measurable and compute them as a power series in λ_0. Two of these, which reduce to m_0 and λ_0 as $\lambda_0 \to 0$, will be termed m and λ.

5. try to eliminate m_0 and λ_0 in favour of m and λ in all physical quantities.

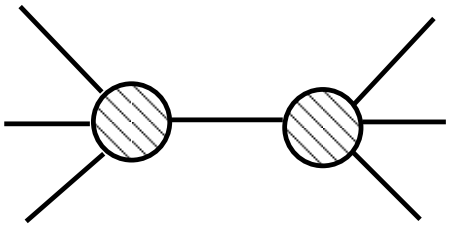

Figure 18: Decomposition of a diagram into one-particle irreducible subdiagrams.

6. if the resultant expressions have a finite limit as the regulator is removed, *i.e.* $\Lambda \to \infty$ or $d \to$ physical dimension, the theory is renormalisable, and we have made sense of it – expressed all measurable observables in terms of a set of others – hopefully a *finite* set otherwise the theory would have no predictive power.

5.1 Analysis of divergences

Divergences occur in loop integrals. Therefore the parts of diagrams which are tree-like are problem-free. We recognise this by breaking each diagram into subdiagrams connected by single lines, as in Fig. 18. Some of these will occur as loop corrections to the external lines, as in Fig. 13. The central pieces left over are

$$\prod_{j=1}^{N} G^{(2)}(p_j)^{-1}\, G^{(N)}(p_1, \ldots, p_N)$$

These are called the truncated N-point functions. However there can still be parts left over which are one-particle reducible (see Fig. 18). We define $\Gamma^{(N)}(p_1, \ldots, p_N)$ to be the one-particle irreducible (1PR) part of the above – by definition, all the diagrams in $\Gamma^{(N)}$ *cannot* be disconnected by breaking a single line. Any divergences must occur within these subdiagrams – if we succeed in making sense of these we make sense of the whole theory. Note that $\Gamma^{(2)}(p) = G^{(2)}(p)^{-1}$.

The integrand in any diagram depends on a number of loop momenta $(k_1, \ldots, k_\ell)$ where ℓ is the number of loops. The most obvious place to look for a UV divergence is where all of these are large and of the same order. Fortunately the degree of these is easy to see just from power counting: if

a diagram has ℓ loops and P propagators, the overall power of momentum is $k^{\ell d}/k^{2P}$, so that if $\delta \equiv \ell d - 2P \geq 0$, the diagram is *primitively* divergent. If $\delta < 0$, on the other hand, we cannot say for sure that it is convergent, since there may be divergences coming from regions where some of the k_j are large and others not. We discuss these later.

The above formula for δ, called the superficial degree of divergence of a diagram, is not very useful because it seems to depend on the details of the diagram. In fact this is not the case, as can be seen by *power-counting*, otherwise known as dimensional analysis. Start with the action

$$S = \int \left[\tfrac{1}{2}(\partial\phi)^2 + \tfrac{1}{2}m_0^2\phi^2 + (\lambda_0/4!)\phi^4\right] d^dx$$

Since S is exponentiated in the path integral, it must be dimensionless in units where $\hbar = 1$. We denote the fact that a quantity X has a momentum dimension of d_X by $[X] = k^{d_X}$. Thus $[p] = k^1$, $[x] = k^{-1}$ and $[S] = k^0$. From this we immediately deduce that $[(\partial\phi)^2] = k^d$ so $[\phi] = k^{(d-2)/2}$. Then $[\lambda_0\phi^4] = [\lambda_0]k^{2(d-2)} = k^d$ so

$$[\lambda_0] = k^{4-d}$$

Now the $G^{(N)}$ in real space are just $\langle\phi(y_1)\dots\phi(y_N)\rangle$ so have dimension $k^{N(d-2)/2}$. Going to momentum space we have to do N integrals $\int d^dx \sim k^{-Nd}$ and divide off an overall momentum-conserving delta function. So

$$[\widetilde{G}^{(N)}] = k^d \cdot k^{-Nd} \cdot k^{N(d-2)/2}$$

Note in particular that $[\widetilde{G}^{(2)}] = k^{-2}$. Thus

$$[\Gamma^{(N)}] = k^{N+d-Nd/2}$$

Finally, these are to be expanded in powers of λ_0

$$\Gamma^{(N)} = \sum_{n=0}^{\infty} \Gamma_n^{(N)} \lambda_0^n$$

and it is the coefficients $\Gamma_n^{(N)}$ which are given by sums of Feynman diagrams. Their dimension gives the superficial degree of divergence of each diagram contributing to $\Gamma^{(N)}$ at $O(\lambda_0^n)$

$$\delta = d + N(1 - d/2) + n(d - 4)$$

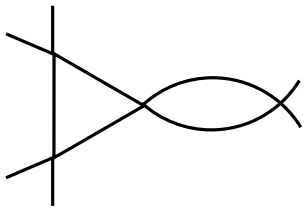

Figure 19: A diagram in $\Gamma^{(6)}$ which is not primitively divergent for $d = 4$ but which contains a divergent subdiagram.

If $\delta \geq 0$, the diagrams are primitively divergent. We distinguish three different cases depending on the value of d:

- $d < 4$: in this case $[\Gamma^{(2)}_n] = k^{2+n(d-4)}$, and all the other δs for $N \geq 4$ are negative. Only a finite number of diagrams, contributing to $\Gamma^{(2)}$ up to order $n \leq 2/(4-d)$, are primitively divergent.
- $d = 4$: $[\Gamma^{(2)}] = k^2$ and $[\Gamma^{(4)}] = k^0$ are primitively divergent to all orders; all other N are not. Note that the coupling constant λ_0 is dimensionless for $d = 4$ – this is called the critical dimension of the theory.
- $d > 4$: all the $\Gamma^{(N)}$ are primitively divergent if evaluated to sufficiently high order.

Note that the critical dimension, and the number of primitively divergent $\Gamma^{(N)}$ at this dimension, depends on the particular theory.

Of course this classification only identifies primitive divergences which occur as all the loop momenta get large. There could, for example, be other divergences for larger N, for example those shown in Fig. 19. However these generally occur in subdiagrams which are of lower order in λ_0 than the whole diagram. Therefore if we implement the renormalisation procedure order-by-order, these divergences will already have been dealt with. Of course, *proving* that this actually works to all orders is quite complicated.

5.2 Mass, field, and coupling constant renormalisation

Let us suppose that d is at, or just below, four, which is the most interesting case for $\lambda\phi^4$ theory. In that case, both $\Gamma_0^{(2)}$ and $\Gamma_0^{(4)}$ have primitive divergences, to arbitrarily high order. We first study those in $\Gamma_0^{(2)}$.

5.2.1 Mass renormalisation

We already evaluated $G_0^{(2)}$ to one loop (see Fig. 7):

$$G_0^{(2)}(p) = \frac{1}{p^2+m_0^2} + \frac{-\lambda_0}{2}\frac{1}{(p^2+m_0^2)^2} I_2 + O(\lambda_0^2)$$

where, in dimensional regularisation,

$$I_2 = \frac{\pi^{d/2}}{(2\pi)^d}\Gamma(1-d/2)m_0^{d-2}$$

or, with a momentum cut-off

$$I_2 = \int_{|p|<\Lambda} \frac{d^dp}{(2\pi)^d}\frac{1}{p^2+m_0^2} \sim \Lambda^{d-2}$$

for $d > 2$, or $\log\Lambda$ for $d = 2$.

Then

$$\Gamma_0^{(2)}(p) = G_0^{(2)}(p)^{-1} = p^2 + m_0^2 - (-\lambda_0/2)I_2 + O(\lambda_0^2)$$

This means that the pole in $G^{(2)}(p)$, or the zero in $\Gamma^{(2)}(p)$, do not occur at $p^2 = -m_0^2$ but at some other value depending on λ_0, m_0 (and Λ). Recall that in Minkowski space this gives the (mass)^2 of the particle, and is therefore physical. We therefore define the *renormalised mass* m in the euclidean theory by

$$\Gamma_0^{(2)}(p^2 = -m^2) = 0$$

Since I_2 is independent of p,

$$m^2 = m_0^2 + (\lambda_0/2)I_2 + O(\lambda_0)^2\,,$$

Note that $m^2 > m_0^2$, so that if we want to consider the massless limit $m^2 \to 0$, then m_0^2 is actually negative. This underlines the fact that it is not a physical parameter.

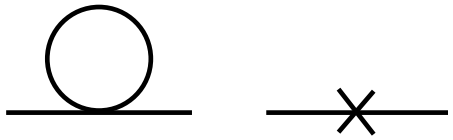

Figure 20: $O(\lambda_0)$ corrections to the 2-pt function, after mass renormalisation. The propagator now carries the physical mass m and the second term represents the counterterm δm^2.

To higher orders the diagrams contributing to $\Gamma^{(2)}(p)$ depend on p, and m is given only implicitly in terms of m_0^2. However, this is not in fact a problem: what we actually do is to write

$$m_0^2 = m^2 + \delta m^2$$

so that, in the lagrangian,

$$\mathcal{L} = \tfrac{1}{2}(\partial\phi)^2 + \tfrac{1}{2}m^2\phi^2 + \tfrac{1}{2}\delta m^2\phi^2 + \cdots$$

We now treat the first two terms as the free theory, so the propagator in Feynman diagrams carries the physical mass m, and we treat the third term (which is $O(\lambda_0)$) as an interaction, denoted in Feynman diagrams by a cross. The $O(\lambda_0)$ contributions to $G_0^{(2)}$ are then as shown in Fig. 20, where, to repeat, the lines correspond to $(p^2+m^2)^{-1}$. The term $\frac{1}{2}\delta m^2\phi^2$ in the lagrangian is called a *counterterm*. Its role is to make sure, order by order in λ_0, that the pole of $G_0^{(2)}(p)$ remains at $p^2 = -m^2$. The beauty of this is that we never actually have to compute δm^2!

5.2.2 Field renormalisation

Exactly at $d = 4$, $\Gamma_0^{(2)}(p)$ diverges quadratically, $\sim \Lambda^2$. This means that there could be subleading terms which are still divergent. In fact if we consider a Taylor expansion about $p^2 = -m^2$

$$\Gamma_0^{(2)}(p) = \Gamma_0^{(2)}(p^2 = -m^2) + (p^2+m^2)\left.\frac{\partial\Gamma_0^{(2)}(p)}{\partial p^2}\right|_{p^2=-m^2} + \cdots,$$

the first term vanishes, but the second term has a superficial degree of divergence $\delta = 0$ and therefore is in general logarithmically divergent in $d = 4$. On the other hand, the higher order terms are superficially finite (which

means that we would find the same logarithmic divergences expanding about some other value of p^2.)

So how should we absorb these divergences? The answer is to remember that $\Gamma_0^{(2)}$ is defined from the 2-point function $\langle\phi_0\phi_0\rangle$, and that $\phi_0(x)$ is just an integration variable in the path integral – by itself it has no physical meaning. So we suppose that there is a *physical* field $\phi(x)$ which is proportional to ϕ_0:

$$\phi(x) = Z_\phi^{-1/2}\phi_0(x)$$

and the physical correlation functions are

$$G^{(N)}(y_1,\ldots,y_N) = \langle\phi(y_1)\ldots\phi(y_N)\rangle = Z_\phi^{-N/2}\langle\phi_0(y_1)\ldots\phi_0(y_N)\rangle$$

Note that this implies that

$$\Gamma^{(N)} = Z_\phi^{N/2}\Gamma_0^{(N)}$$

The factor $Z_\phi^{-1/2}$ is called the field renormalisation constant (or wave-function renormalisation in some texts.) It is fixed by requiring

$$\left.\frac{\partial\Gamma^{(2)}(p)}{\partial p^2}\right|_{p^2=-m^2} = 1$$

Equivalently

$$Z_\phi^{-1} = \left.\frac{\partial\Gamma_0^{(2)}(p)}{\partial p^2}\right|_{p^2=-m^2}$$

Note that $Z_\phi = 1 + 0(\lambda_0^2)$ in this theory, because I_2 does not depend on the external momentum. This is not the case for a general theory.

5.2.3 Coupling constant renormalisation

In this theory $\Gamma_0^{(4)}$ is superficially logarithmically divergent in $d = 4$ to all orders. Since to lowest order it is just given by $-\lambda_0$, this suggest that we define the renormalised coupling constant λ in terms of $\Gamma^{(4)}$. As we shall show later, in Minkowski space $\Gamma^{(4)}(p_1, p_2, -p_3, -p_4)$ gives the scattering amplitude for particle of 4-momenta (p_1, p_2) to scatter into (p_3, p_4), so it

is certainly physically measurable. However, we want to define λ as a number, not a function, which means we have to specify particular values for the momenta. One choice which makes sense in the scattering context is to choose its value at zero relative 3-momenta, *i.e.*

$$i\lambda \equiv \Gamma^{(4)}((m,\mathbf{0}),(m,\mathbf{0}),(-m,\mathbf{0}),(-m,\mathbf{0}))$$

in Minkowski space.

5.2.4 Renormalisation schemes

We have described one particular way of defining the renormalised theory, motivated by the particle-scattering application of the theory. This is often termed mass-shell renormalisation. However, since primitive divergences are more or less independent of the external momenta, we may choose other schemes which are equally valid and often easier to compute with, and physically relevant for other applications. Renormalised correlation functions in different schemes are related by transformations which are finite as the regulator is removed.

An example is zero-momentum normalisation: in euclidean space

$$m^2 = \Gamma^{(2)}(p=0) \qquad \left.\frac{\partial\Gamma^{(2)}(p)}{\partial p^2}\right|_{p=0} = 1 \qquad \lambda = -\Gamma^{(4)}(p_1 = \ldots = p_4 = 0)$$

Let us compute the renormalised 4-point function to one loop for $d = 4$ in this scheme.

$$\Gamma^{(4)}(p_1,\ldots,p_4) = Z_\phi^2\left(-\lambda_0 + \tfrac{1}{2}\lambda_0^2\left[I_4(p_1+p_2)+I_4(p_1+p_3)+I_4(p_1+p_4)\right] + O(\lambda_0^3)\right)$$

where, as we computed earlier

$$I_4(q) = \frac{\pi^{d/2}}{(2\pi)^d}\Gamma(2-d/2)\int_0^1 [x(1-x)q^2+m^2]^{d/2-2}dx$$

Since $Z_\phi = 1 + O(\lambda_0^2)$ in this theory, we set it $= 1$ to the order required, so

$$\lambda = \lambda_0 - \tfrac{3}{2}\lambda_0^2\frac{\pi^{d/2}}{(2\pi)^d}\Gamma(2-d/2)m^{d-4} + O(\lambda_0^3)$$

Solving for λ_0 in terms of λ

$$\lambda_0 = \lambda + \tfrac{3}{2}\lambda^2 \frac{\pi^{d/2}}{(2\pi)^d}\Gamma(2-d/2)m^{d-4} + O(\lambda^3)$$

Inserting this into the expansion for $\Gamma^{(4)}$

$$\begin{aligned}\Gamma^{(4)}(p_1,\ldots,p_4) &= -\lambda + \tfrac{1}{2}\lambda^2 \frac{\pi^{d/2}}{(2\pi)^d}\Gamma(2-d/2) \\ &\quad \times \left(\int_0^1 ([x(1-x)(p_1+p_2)^2+m^2]^{d/2-2} - m^{d-4}]dx + \text{perms}\right)\end{aligned}$$

Removing the regulator, that is letting $d \to 4$, and recalling that $\Gamma(2 - d/2) \sim 2/(4-d)$ in that limit,

$$\Gamma^{(4)}(p_1,\ldots,p_4) = -\lambda - \frac{\lambda^2}{32\pi^2}\left(\int_0^1 \log\left(\frac{x(1-x)(p_1+p_2)^2+m^2}{m^2}\right)dx + \text{perms}\right)$$

As advertised, the renormalised correlation functions are finite when expressed in terms of the renormalised parameters. Note however, that the logarithm in the integrand is not anything we could have got out of the elementary Feynman rules – even the sign of the $O(\lambda^2)$ term is different!

This example points the way to another, even simpler, renormalisation scheme. We can write, in general,

$$\Gamma^{(4)} = Z_\lambda^{-1}\Gamma_0^{(4)} \tag{11}$$

where $Z_\lambda = 1 + O(\lambda) = 1 + O(\lambda_0)$. The important property of this $O(\lambda)$ term is that it has a pole as $d \to 4$ which, perturbatively, cancels the divergence in $\Gamma^{(4)}$. In fact we see from the above that

$$Z_\lambda = 1 - \frac{\lambda_0}{m^\epsilon}\left(\frac{1}{16\pi^2\epsilon} + \ldots\right) + O(\lambda_0)^2$$

where $\epsilon \equiv 4-d$ (not to be confused with the ϵ of the $i\epsilon$ prescription) and the omitted terms in the parentheses are finite as $\epsilon \to 0$. *Minimal subtraction* renormalisation amounts to neglecting this finite term altogether, that is, we *define* the renormalised $\lambda = Z_\lambda\lambda_0$, with the condition that Z_λ contains *only* pole terms in ϵ with coefficients fixed by the requirement that the lhs

of (11) is finite as $\epsilon \to 0$. The field renormalisation is defined similarly, and mass renormalisation simply does not show up in this scheme, since it corresponds to poles at $d = 2$.

However neither of the above schemes work in the case where the renormalised mass $m = 0$, because of infrared (IR) divergences. (Think, for example about the 1-loop coupling constant correction, which would involve the integral $\int(d^d p/p^4)$ - UV divergent for $d \geq 4$ and IR divergent for $d \leq 4$.) Massless theories are important because gauge theories are massless, at least perturbatively, and in critical behaviour being massless is the same as being at the critical point. So we need to know how to renormalise them. The way out is to introduce an extra parameter μ with the dimension of mass, and renormalise at values of the external momenta proportional to μ

$$m^2 = \Gamma^{(2)}(p = 0) = 0 \quad \text{but} \quad \left.\frac{\partial\Gamma^{(2)}(p)}{\partial p^2}\right|_{p^2=\mu^2} = 1, \qquad \lambda = -\Gamma^{(4)}(p_j \sim \mu)$$

It will turn out that the necessity to introduce such a scale has dramatic consequences.

For $d < 4$ the IR divergences of the massless theory are even worse. In fact the only way to define the renormalised massless theory perturbatively for $d < 4$ is in a double expansion in λ and $4 - d$.

We are now ready to make the *statement of renormalisability* for $\lambda\phi^4$ theory. Starting with the regularised bare theory, parameterised by m_0 and λ_0, if we make:

- field renormalisation, $\phi = Z_\phi^{-1/2}\phi_0$ such that, *e.g.* $\partial\Gamma^{(2)}(p)/\partial p^2|_{p=0} = 1$;
- mass renormalisation, *e.g.* $m^2 = \Gamma^{(2)}(p = 0)$;
- coupling constant renormalisation, *e.g.* $\lambda = -\Gamma^{(4)}(p_j = 0)$.

Then, for $d \leq 4$, all renormalised N-point functions $\widetilde{G}^{(N)}(p_1, \ldots, p_N)$ have a finite limit as the regulator is removed, *when expressed in terms of the renormalised mass m and coupling constant λ.*

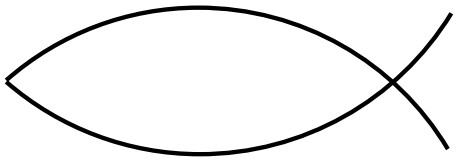

Figure 21: One loop correction to the composite field ϕ^2.

Note that this guarantees the finiteness of the correlation functions in position space

$$\langle \phi(y_1) \dots \phi(y_N) \rangle = \int \prod_{j=1}^{N} e^{ip_j y_j} d^d p_j \widetilde{G}^{(N)}(p_1, \dots, p_N)(2\pi)^d \delta^{(d)}(\sum_j p_j)$$

only if the points $\{y_j\}$ do not coincide. In this case the additional momentum integrals are damped by phase oscillations of factors like $e^{i(y_j - y_{j'})p_j}$. When some of the $\{y_j\}$ do coincide there are still divergences. For example, to one loop

$$\langle \phi_0(y_1)^2 \phi_0(y_2) \phi_0(y_3) \rangle$$

is given by the diagram in Fig. 21, which is logarithmically divergent and is not made finite by field or coupling constant renormalisation. This is an example of a *composite field*. It requires further renormalisation

$$\phi^2(y) = Z_{\phi^2}^{-1} \, \phi_0(y)^2$$

where $Z_{\phi^2} \neq 1$ is fixed by requiring that the Fourier transform of $\langle \phi^2(y_1) \phi(y_2) \phi(y_3) \rangle$ is finite at a suitably chosen normalisation point. One way to state this is that, in renormalised QFT,

$$\phi^2(x) \neq \phi(x) \cdot \phi(x) \qquad (!!)$$

Similar additional renormalisations are in principle required for all composite fields, *e.g.* ϕ^3, $(\partial\phi)^2$, and so on. An important exception are conserved currents J_μ because their integrals give physical charges. However, some currents are *anomalous*: while conserved in the classical theory their quantum versions are not.

5.2.5 Renormalisation in other theories

The process of renormalisation in scalar theories with other interaction lagrangians follows the same steps as for $\lambda\phi^4$ theory (gauge theories are

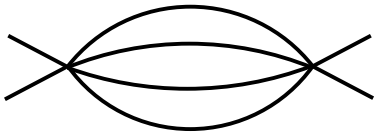

Figure 22: Loop correction to $\Gamma^{(4)}$ in ϕ^6 theory which is linearly divergent in $d = 3$ and requires subtraction.

more complicated.)

- do power counting to identify the primitively divergent $\Gamma^{(N)}$;
- identify the critical dimension d_c for which the coupling constant becomes dimensionless: for $d < d_c$ only a finite number of $\Gamma^{(N)}$ are primitively divergent and only to a finite order, while for $d = d_c$ only a finite number diverge, but to all orders;
- at or just below d_c, make finite the power law diverges in by subtractions, as for the mass in ϕ^4 theory;
- remaining logarithmic divergences are taken into account by field and coupling constant renormalisation.

As an example, consider $\kappa\phi^6$ theory. The previous result

$$[\Gamma^{(N)}] = k^{N+d-Nd/2}$$

still holds, but now we see that

$$[\kappa] = [\Gamma^{(6)}] = k^{6-2d}$$

so $d_c = 3$. At d_c, $[\Gamma^{(2)}] = k^2$, $[\Gamma^{(4)}] = k^1$ and $[\Gamma^{(6)}] = k^0$. This means that this theory requires mass, field and coupling constant renormalisation as before, but also a counterterm coupling to ϕ^4 in the lagrangian. This is because even if this term was not present in the bare theory, it gets generated by loop corrections, for example Fig. 22. Another way to state this is, in order to make the renormalised ϕ^4 coupling λ vanish, so we are talking about ϕ^6 rather than ϕ^4 theory, then the bare ϕ^4 coupling $\lambda_0 \neq 0$ and in fact is linearly divergent.

6 Renormalisation Group

We often want to understand the behaviour of correlation functions in momentum space as $p \to \infty$ (their UV, or high-energy, behaviour), or, in the massless theory, as $p \to 0$ (their IR behaviour). This is the object of the *renormalisation group* (RG). Note that, despite its name, this is not a group in the mathematical sense. Also, some students may be confused having seen the so-called real-space RG in the context of lattice statistical mechanics models. The two sets of ideas are related, but we shall not discuss this here (see separate notes on the web page.)

One naive way to understand the p-dependence might be through dimensional analysis. For example, we know that $[G^{(2)}] = k^{-2}$ so we might suppose that $G^{(2)}(p) \sim 1/p^2$ when p is larger than all the other quantities in the theory with the dimensions of mass. However, this is ignoring the fact that the renormalised theory is defined through the limit $\Lambda \to \infty$ of a cut-off theory, and so p is never larger than this implicit scale Λ. Thus all we can really assert on the basis of dimensional analysis is

$$G_0^{(2)}(p) \sim \frac{1}{p^2} F(p^2/\Lambda^2)$$

where F is some presently unknown function. Equivalently, in the renormalised massless theory (which we shall mainly consider)

$$G^{(2)}(p) \sim \frac{1}{p^2} \widetilde{F}(p^2/\mu^2) \tag{12}$$

where μ is the renormalisation scale discussed in the last section and $\widetilde{F}$ is another unknown function.

We note from (12) that instead of considering the p-dependence of $G^{(2)}$ we may equally well study its dependence on μ. The renormalisation group (RG) is a way of doing this.

6.1 Callan-Symanzik equation

In what follows it is useful to define the dimensionless renormalised coupling constant

$$g \equiv \lambda \mu^{-\epsilon}$$

where $\epsilon = 4 - d$ as before. We shall assume that the renormalised mass $m = 0$, only later seeing how things change in the massive case.

The statement of renormalisability says that

$$\Gamma^{(N)}(\{p_j\}, g, \mu) = Z_\phi^{N/2}(\lambda_0, \mu)\ \Gamma_0^{(N)}(\{p_j\}, \lambda_0)$$

has a finite limit as the regulator (in this case dimensional, since we haven't included Λ in the argument of $\Gamma_0^{(N)}$) is removed. In writing this we have been careful to exhibit of which variables the various quantities are considered to be functions. Note that we could equally (and indeed will) consider Z_ϕ to be a function of g and μ.

Now we exploit the simple but powerful fact that the bare vertex functions $\Gamma_0^{(N)}$ do not know anything about the renormalisation scale μ. Thus

$$\mu \frac{\partial}{\partial \mu} \Gamma_0^{(N)}(\{p_j\}, \lambda_0)\Big|_{\lambda_0} = 0$$

where the derivative is taken keeping the bare coupling λ_0 fixed (the momenta $\{p_j\}$ are held fixed until further notice). In terms of the renormalised vertex functions this seems less of a tautology:

$$\mu \frac{\partial}{\partial \mu} \left(Z_\phi^{-N/2}(g, \mu)\, \Gamma^{(N)}(\{p_j\}, g, \mu) \right)\Big|_{\lambda_0} = 0$$

Using the chain rule

$$\left(\mu \frac{\partial}{\partial \mu} + \mu \left.\frac{\partial g}{\partial \mu}\right|_{\lambda_0} \frac{\partial}{\partial g} - \frac{N}{2} Z_\phi^{-1} \mu \left.\frac{\partial Z_\phi}{\partial \mu}\right|_{\lambda_0} \right) \Gamma^{(N)}(\{p_j\}, g, \mu) = 0$$

where the first derivative now acts only on the explicit μ dependence of $\Gamma^{(N)}$. Introducing the functions

$$\beta(g) \equiv \mu \left.\frac{\partial g}{\partial \mu}\right|_{\lambda_0}, \qquad \gamma_\phi(g) \equiv \mu \left.\frac{\partial \log Z_\phi}{\partial \mu}\right|_{\lambda_0}$$

this gives the massless version of the *Callan-Symanzik equation*

$$\left(\mu\frac{\partial}{\partial\mu}+\beta(g)\frac{\partial}{\partial g}-\frac{N}{2}\gamma_\phi(g)\right)\Gamma^{(N)}(\{p_j\},g,\mu)=0 \tag{13}$$

This equation says that we can trade the dependence of the vertex functions on μ for their dependence on g.

6.2 Renormalisation group flows

If we now use dimensional analysis (12) for the case $N = 2$ (the argument generalises straightforwardly to larger N as long as we scale all the momenta proportionally) we see that $\Gamma^{(2)}$ satisfies the Euler equation

$$\left(\mu\frac{\partial}{\partial\mu}+p\frac{\partial}{\partial p}-2\right)\Gamma^{(2)}(p,g,\mu)=0$$

Subtracting this from the C-S equation (13)

$$\left(p\frac{\partial}{\partial p}-\beta(g)\frac{\partial}{\partial g}-(2-\gamma_\phi(g))\right)\Gamma^{(2)}(p,g,\mu)=0 \tag{14}$$

This means that we can trade dependence on p, which is what we are after, for dependence on g. In particular:

- if $\beta(g) > 0$ then $p\uparrow \Leftrightarrow g\uparrow$ and $p\downarrow \Leftrightarrow g\downarrow$
- if $\beta(g) < 0$ then $p\uparrow \Leftrightarrow g\downarrow$ and $p\downarrow \Leftrightarrow g\uparrow$

The really interesting case is when $\beta(g)$ has a zero at some value $g = g^*$. Then if $\beta(g)$ changes sign from < 0 to > 0 as $g\uparrow$ through g^* (as on the left in Fig. 23), then as $p \to 0$, $g \to g^*$. This is called an IR stable zero. Conversely if $\beta(g)$ changes sign from > 0 to < 0 as $g\uparrow$ through g^* (as on the right in Fig. 23), then as $p \to \infty$, $g \to g^*$. This is called a UV stable zero.

This is clearer in the explicit solution of the PDE (14): define the *running coupling* $g(p)$ as the solution of the ordinary differential equation

$$p\frac{\partial}{\partial p}g(p)=\beta(g(p))$$

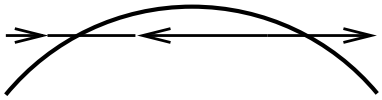

Figure 23: A beta-function with an IR stable zero (on the left) and a UV stable zero (right). The RG flows as $p \to 0$ are indicated.

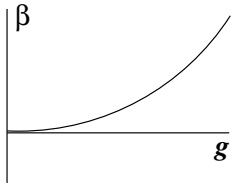

Figure 24: Form of the beta-function in ϕ^4 theory for $d = 4$.

with the initial condition that $g(\mu) = g$. This is called an *RG flow* equation. It tells us how $g(p)$ flows as we change p. The solution of (14) is then

$$\Gamma^{(2)}(p, g, \mu) = \exp\left(\int_g^{g(p)} \frac{2 - \gamma_\phi(g')}{\beta(g')} dg'\right) \Gamma^{(2)}(\mu, g(p), \mu) \qquad (15)$$

This shows how the dependence on p in the lhs gets traded for the dependence on $g(p)$ on the rhs.

Different renormalisable QFTs have different beta-functions and therefore qualitatively different UV and IR behaviours

For example, the beta function for $\lambda\phi^4$ theory, which we shall compute shortly to one loop, has the forms of Fig. 24 for $d = 4$ and Fig. 25 for $d < 4$. In the first case there is a single IR stable zero at $g = 0$. In the second, this moves to $g = g^* > 0$ and the zero at $g = 0$ becomes UV stable. For QCD in $d = 4$ the beta function looks like that in Fig. 26: there is a UV stable zero at $g = 0$. This is *asymptotic freedom*: the UV behaviour of

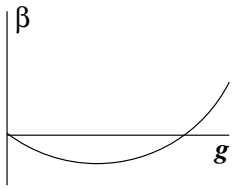

Figure 25: Form of the beta-function in ϕ^4 theory for $d < 4$.

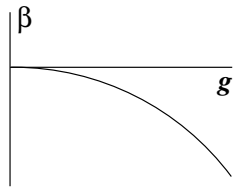

Figure 26: Form of the beta-function in QCD for $d = 4$.

QCD can be computed perturbatively in $g(p)$. Conversely, as $p \to 0$ (large distances), $g(p) \to \infty$, which is consistent with the idea of *confinement*.

If the theory has a UV (IR) stable fixed point at $g = g^*$ then we can get the leading behaviour of the correlation functions as $p \to \infty$ ($p \to 0$) by simply setting $g = g^*$ in (14). Thus

$$\left(p\frac{\partial}{\partial p} - (2 - \gamma^*)\right)\Gamma^{(2)}(p) = 0$$

where $\gamma^*_\phi = \gamma_\phi(g^*)$. This has the simple solution

$$\Gamma^{(2)}(p) \propto p^{2-\gamma^*_\phi}$$

Thus we see that (12) becomes

$$G^{(2)}(p, \Lambda) \sim \frac{1}{p^2}\left(\frac{p}{\mu}\right)^{\gamma^*_\phi}$$

or, in position space

$$\langle\phi(x)\phi(0)\rangle \sim \frac{1}{|x|^{d-2+\gamma^*_\phi}}$$

One way to think of this is that the field $\phi(x)$, instead of having its canonical dimension $k^{(d-2)/2}$, instead has dimension

$$\tfrac{1}{2}(d-2) + \tfrac{1}{2}\gamma^*_\phi$$

This is sometimes called the *scaling dimension* of ϕ. The difference $\frac{1}{2}\gamma^*_\phi$ between this and the canonical dimension is called the *anomalous dimension*.

6.3 One-loop computation in $\lambda\phi^4$ theory

In minimal subtraction, the dimensionless renormalised coupling is

$$g = \lambda\mu^{-\epsilon} = \mu^{-\epsilon}\left(\lambda_0 - \frac{3}{16\pi^2}\frac{1}{\epsilon}\lambda_0^2\mu^{-\epsilon} + O(\lambda_0^2)\right)$$

Then

$$\begin{aligned}\beta(g) = \mu\left.\frac{\partial g}{\partial\mu}\right|_{\lambda_0} &= -\epsilon g + \mu^{-\epsilon}\left(\frac{3}{16\pi^2}\lambda_0^2\mu^{-\epsilon} + O(\lambda_0^3)\right)\\ &= -\epsilon g + \frac{3}{16\pi^2}(\lambda_0\mu^{-\epsilon})^2 + O(\lambda_0^3)\\ &= -\epsilon g + \frac{3}{16\pi^2}g^2 + O(g^3)\end{aligned}$$

For $\epsilon = 0$ we see that $\beta(g)$ has the form shown in Fig. 24 with an IR stable zero at $g = 0$, while for ϵ small and > 0, it looks like Fig. 25, and there is an IR stable zero at

$$g^* = \frac{16\pi^2\epsilon}{3} + O(\epsilon^2)$$

To find the higher order terms in ϵ we would have to carry the calculation of $\Gamma^{(2)}$ and $\Gamma^{(4)}$ to higher loop order.

6.3.1 Calculation of γ_ϕ

As we have already observed, in this theory $Z_\phi = 1$ to one-loop order, so we have to go to two loops to get a nontrivial result for γ_ϕ, which involves the last diagram in Fig. 9.

$$Z_\phi^{-1} = \left.\frac{\partial\Gamma_0^{(2)}(p)}{\partial p^2}\right|_{p^2=\mu^2} = 1 - \frac{\lambda_0^2}{3!}\frac{\partial}{\partial p^2}\int\left.\frac{(d^dk_1 d^dk_2/(2\pi)^{2d})}{k_1^2k_2^2(p-k_1-k_2)^2}\right|_{p^2=\mu^2} + O(\lambda_0^3)$$

The integral can be done by the methods described earlier and we get

$$Z_\phi^{-1} = 1 + \frac{1}{12}\frac{1}{\epsilon}\frac{\lambda_0^2\mu^{-2\epsilon}}{(16\pi^2)^2} + O(\lambda_0^3)$$

so

$$\gamma_\phi = \mu\frac{\partial}{\partial\mu}\log Z_\phi = \frac{1}{6}\left(\frac{g}{16\pi^2}\right)^2 + O(g^3)$$

and, at the IR stable fixed point $g = g^*$,

$$\gamma_\phi^* = \frac{\epsilon^2}{54} + O(\epsilon^3)$$

(Note that we do not need the $O(\epsilon^2)$ term in g^* to get this.)

For $d = 4$, the IR stable zero is at $g = 0$ and so $\gamma_\phi^* = 0$. This does not mean that $G^{(2)}(p) \sim 1/p^2$ as $p \to 0$, however, since in this case $g(p) \to 0$ very slowly. In fact there are calculable logarithmic factors in $G^{(2)}(p)$, which can be found using (15).

6.3.2 Anomalous dimension of $\phi^2(x)$

As discussed earlier, there are additional divergences in correlation functions of products of operators at the same point. As an example, consider

$$G_0^{(2,1)}(p_1, p_2) \equiv \int d^d y_1 d^d y_2 e^{i(p_1 y_1 + p_2 y_2)} \langle \phi_0(0)^2 \phi_0(y_1) \phi_0(y_2) \rangle$$

(The 1 in $(2, 1)$ means that there is one insertion of ϕ^2 in the 2-point function.) In order to define the corresponding vertex function we should truncate the external legs. Thus

$$\Gamma_0^{(2,1)}(p_1, p_2) = \frac{G_0^{(2,1)}(p_1, p_2)}{G_0^{(2)}(p_1) G_0^{(2)}(p_2)}$$

We absorb the divergences in ϕ^2 through the renormalisation constant Z_{ϕ^2}:

$$\phi^2(x) = Z_{\phi^2}^{-1} \phi_0(x)^2$$

Hence the renormalised vertex function is related to the bare one by

$$\Gamma^{(2,1)}(p_1, p_2) = \frac{G^{(2,1)}(p_1, p_2)}{G^{(2)}(p_1) G^{(2)}(p_2)} = Z_{\phi^2}^{-1} Z_\phi \Gamma_0^{(2,1)}(p_1, p_2)$$

where Z_{ϕ^2} is fixed, for example, $\Gamma^{(2,1)}(p_1, p_2)|_{(p_1+p_2)^2=\mu^2} = 1$, or by using minimal substraction.

The C-S equation follows as before from the fact that this is independent of μ:

$$\left(\mu \frac{\partial}{\partial \mu} + \beta(g) \frac{\partial}{\partial g} - \gamma_\phi(g) + \gamma_{\phi^2}(g) \right) \Gamma^{(2,1)}(p_1, p_2, \mu) = 0$$

where $\gamma_{\phi^2} \equiv (\mu\partial/\partial\mu)\log Z_{\phi^2}|_{\lambda_0}$. Using the fact that $\Gamma^{(2,1)}$ is dimensionless, we can trade $\mu\partial/\partial\mu$ for $-p\partial/\partial p$. Going to the fixed point $g = g^*$ we therefore find that

$$\Gamma^{(2,1)}(p_1, p_2) \sim p^{\gamma^*_{\phi^2} - \gamma^*_\phi} \tag{16}$$

as $p_1 \sim p_2 \sim p \to 0$. Thus we can think of $\phi^2/(\phi\cdot\phi)$ as having anomalous dimension $\gamma^*_{\phi^2}$.

The diagram for $\Gamma^{(2,1)}_0(p_1, p_2)$ to one loop is that shown in Fig. (21), with the external lines truncated. The loop integral is the same as that for the 1-loop coupling constant renormalisation, so

$$Z_{\phi^2} = 1 - \frac{1}{32\pi^2}\frac{2}{\epsilon}\lambda_0\mu^{-\epsilon} + O(\lambda_0^2)$$

so

$$\gamma^*_{\phi^2} = \mu\frac{\partial}{\partial\mu}\log Z_{\phi^2} = \frac{\lambda_0\mu^{-\epsilon}}{16\pi^2} + O(\lambda_0^2) = \frac{g}{16\pi^2} + O(g^2)$$

At the fixed point $g = 16\pi^2\epsilon/3 + O(\epsilon^2)$ we therefore have

$$\gamma^*_{\phi^2} = \frac{\epsilon}{3} + O(\epsilon^2)\,.$$

6.4 Application to critical behaviour in statistical mechanics

We have remarked earlier that the euclidean path integral

$$\int [d\phi] e^{-S_E[\phi]/\hbar}$$

is formally very similar to the partition function in classical statistical mechanics

$$\mathrm{Tr}\, e^{-E/k_B T}$$

In fact we can make the connection more explicit. It turns out that many aspects of the behaviour of renormalised $\lambda\phi^4$ theory in the limit of small renormalised mass are the same as those of the d-dimensional Ising model near its critical point.

The Ising model is defined on a lattice. At each point x of the lattice resides a 'spin' $s(x)$ taking the values ± 1. The interaction energy between

these spins is

$$E[s] = -\tfrac{1}{2}\sum_{x,x'} J(x-x')s(x)s(x')$$

where $J > 0$ is a short-ranged function. At high temperatures the system is paramagnetic and the 2-point correlation function

$$\langle s(x_1)s(x_2)\rangle = \frac{1}{Z}\,\mathrm{Tr}\, s(x_1)s(x_2)e^{-E[s]/kT}$$

decays as $\sim e^{-|x_1-x_2|/\xi}$, where ξ is the correlation length. The Fourier transform of this is $1/(p^2+\xi^{-2})$, and so we can think of ξ^{-1} as being similar to the renormalised mass m. At the temperature is reduced, ξ increases, and, for $d > 1$, there is a continuous phase transition at some temperature T_c. Below T_c the system orders: that is, if we take the limit as $H \to 0$ of an applied magnetic field, the 1-point function $\langle s(x)\rangle$ is non-zero. The main point is that at T_c, $m = 0$.

However, the Ising model is not even a lattice field theory since the spins take only the values ± 1. On way to map it to a field theory is the *Hubbard-Stratonovich transformation*. Since this is an important tool in other applications, we shall discuss it in detail. The Ising model partition function is

$$Z[H] = \sum_{s(x)} e^{\frac{1}{2}\sum_{x,x'} s(x)K(x-x')s(x') + \sum_x H(x)s(x)}$$

where $K = J/kT$ and we have added a source $H(x)$ (a magnetic field) in order to compute correlation functions through differentiation.

Now use a gaussian integral (similar to the computation of the propagator)

$$e^{\frac{1}{2}\sum_{x,x'} s(x)K(x-x')s(x')} \propto \int \prod_x d\phi(x) e^{-\frac{1}{2}\sum_{x,x'}\phi(x)K^{-1}(x-x')\phi(x') + \sum_x \phi(x)s(x)}$$

where $K^{-1}(x-x')$ is the matrix inverse of $K(x-x')$ with x and x' labelling the rows and columns. We can now do the sum over $s(x)$ at each site:

$$\sum_{s(x)=\pm 1} e^{\big(\phi(x)+H(x)\big)s(x)} = 2\cosh\left(\phi(x)+H(x)\right)$$

Thus, apart from an unimportant constant, the partition function is that of a euclidean lattice field theory with action

$$S = \frac{1}{2}\sum_{x,x'} \phi(x) K^{-1}(x-x')\phi(x') - \sum_x \log\cosh\left(\phi(x) + H(x)\right)$$

So far this is exact. Now we take the naive continuum limit and make a gradient expansion of the first term:

$$\begin{aligned}
&\tfrac{1}{2}\sum_{x,x'} \phi(x) K^{-1}(x-x')\phi(x') = \tfrac{1}{2}\int \frac{d^dp}{(2\pi)^d} \tilde\phi(p)^* \widetilde K(p)^{-1} \tilde\phi(p) \\
&\approx \int \frac{d^dx}{a^d} \widetilde K(0)^{-1} \left(\phi(x)^2 + a^2R^2(\partial\phi(x))^2 + \cdots\right)
\end{aligned}$$

where $\widetilde K(p) = \sum_x e^{ipx} K(x) = \widetilde K(0)(1 - R^2p^2 + O(p^4))$, so that $R^2 = \sum_x x^2 K(x) / \sum_x K(x)$. ($R$ can be thought of as the range of the interaction.) Similarly we expand the second term in powers of $\phi(x)$ and $H(x)$. After rescaling ϕ so that

$$\frac{R^2 a^{-d}}{\widetilde K(0)} (\partial\phi)^2 \to (\partial\phi)^2$$

we end up with the usual ϕ^4 action, plus corrections.

$$S = \int \left[\tfrac{1}{2}(\partial\phi)^2 + \tfrac{1}{2}m_0^2 + \tfrac{\lambda_0}{4!}\phi^4 + J(x)\phi(x) + \cdots\right] d^dx \tag{17}$$

where

$$m_0^2 \propto \widetilde K(0)^{-1} - 1\,, \qquad \lambda_0 \propto a^d/R^4\,, \qquad J(x) \propto H(x) + O(H^3)$$

If we ignore the corrections for the time being, we see first that the bare mass m_0 vanishes when $\widetilde K(0) = \sum_x J(x)/kT = 1$. This corresponds to the mean field approximation to the critical temperature $T_{MF} = k^{-1}\sum_x J(x)$. The actual critical point occurs where the renormalised mass m vanishes, and, since $m_0^2 < m^2$, we see that the true $T_c < T_{MF}$. Second, the bare coupling λ_0 is small if the range of the interaction R is large. Since the actual expansion parameter is $\lambda_0 m^{d-4}$ we see that for $d < 4$, if the range of interaction R is large the corrections to mean field theory are small except in a narrow window $|m| < O(R^{-4/(4-d)})$. This is called the Ginzburg criterion.

However we have not yet justified the neglect of the higher order terms in (17). These contain powers of ϕ higher than 4, and/or powers of ∂_x higher than 2. The important point is that, after rescaling the field, they all enter the lagrangian with *positive* powers of a, near $d = 4$. If we consider such a term it will have the form $\kappa \int \mathcal{O} d^d x$ where $\mathcal{O}$ is some polynomial in ϕ and ∂, and $[\kappa] = k^{-\delta}$ with $\delta > 0$ (for d just below 4). This means that if we now do perturbation theory in κ, the dimensionless expansion parameter is $h = \kappa\mu^\delta$, and the corresponding renormalisation group function has the form

$$\beta_h(h) = \delta \cdot h + O(h^2)$$

This means that $h = 0$, or $\kappa = 0$ is an IR stable fixed point. For small enough κ, the the IR behaviour is the same as if we took $\kappa = 0$. This is not to say that there might not be some other behaviour for larger values of κ. However, for the values corresponding to the nearest neighbour spin-$\frac{1}{2}$ Ising model, there is ample numerical evidence that it has the same critical behaviour as ϕ^4 field theory. The phenomenon that two very different looking models have the same scaling behaviour is called *universality*.

However, note that in statistical physics the UV cut-off $\Lambda \sim a^{-1}$ is finite, and the physical correlation functions are the bare ones. However, since, as far as their x-dependence goes, these are proportional to those of the renormalised theory, we can take over results from the latter.

For example, at $T = T_c$ ($m = 0$) we have

$$\langle s(x)s(0)\rangle \propto \langle \phi(x)\phi(0)\rangle \sim \frac{1}{|x|^{d-2+\gamma^*_\phi}}$$

In the statistical physics literature, the exponent γ^*_ϕ is denoted by η.

In statistical physics, we are interested in how the behaviour close to the critical point depends on $T - T_c = \delta T \propto m_0^2 - m_{0c}^2$. In particular, we want to know how the correlation length ξ diverges as $T \to T_c$, that is, how m depends on $m_0^2 - m_{0c}^2$. This defines the critical exponent ν:

$$\xi^{-1} = m \sim (\delta T)^\nu \propto (m_0^2 - m_{0c}^2)^\nu$$

The exponent ν is related to the scaling dimension of the composite operator ϕ^2. The simplest, if not the most rigorous, way to see this is through

a scaling argument: the change in the action is

$$\delta S \propto (T - T_c) \int \phi_0^2 d^d x$$

Since $[\phi_0^2] = k^{d-2+\gamma^*_{\phi^2}}$, and δS is dimensionless,

$$[T - T_c] = k^{2-\gamma^*_{\phi^2}} = [\xi]^{-2+\gamma^*_{\phi^2}}$$

so that

$$\nu = \frac{1}{2 - \gamma^*_{\phi^2}}$$

and, using our 1-loop result for $\gamma^*_{\phi^2}$,

$$\nu = \frac{1}{2 - \frac{1}{3}\epsilon + O(\epsilon^2)} = \frac{1}{2} + \frac{\epsilon}{12} + O(\epsilon^2)\,.$$

6.5 Large N

The ϕ^4 theory we have been discussing is a special case of a more general theory where the field $\phi_a(x)$ has N components, $a = 1, \ldots, N$. The most interesting case is when the lagrangian has O(N) symmetry:

$$\mathcal{L} = \tfrac{1}{2} \sum_{a=1}^{N} (\partial \phi_a)^2 + \tfrac{1}{2} m_0^2 \sum_{a=1}^{N} \phi_a^2 + \tfrac{1}{8} \lambda_0 (\sum_{a=1}^{N} \phi_a^2)^2$$

In critical behaviour, $N = 1$ corresponds to the Ising magnets, $N = 2$ to XY magnets (and also a complex scalar field representing the macroscopic wave-function in superfluids and superconductors), $N = 3$ to Heisenberg magnets, and so on. It turns out that the $N \to \infty$ limit is soluble and nontrivial, and also is the starting point of a systematic $1/N$ expansion.

There are several ways of analysing this limit but we shall restrict ourselves to the one closest to the earlier parts of this course. We must first understand how to draw Feynman diagrams for general N. Write the interaction part of the action using (yet again!) a gaussian transformation

$$e^{-(\lambda_0/8) \int \left(\sum_{a=1}^{N} \phi_a^2 \right)^2 d^d x} \propto \int [du(x)] e^{\int [-\frac{1}{2} u^2 + (i \lambda_0^{1/2}/2) u \sum_a \phi_a^2] d^d x}$$

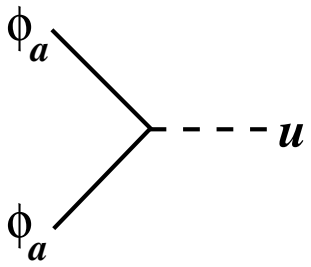

Figure 27: Vertex between two ϕ_a fields and the auxiliary field u.

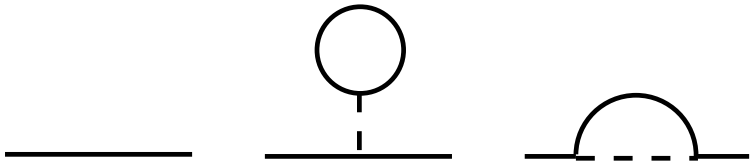

Figure 28: First few diagrams for the 2-point function. The second is $O(N)$ relative to the third.

We now have two types of field, ϕ_a, with a bare propagator

$$\langle \phi_a(x_1)\phi_b(x_2)\rangle = \delta_{ab}\,\Delta(x_1 - x_2)$$

and u, with a propagator, in p-space, equal to 1. Denoting this by a dashed line, the interaction vertex is as shown in Fig. 27, and takes the value $i\lambda_0$. The first few diagrams for the 2-point function are shown in Fig. 28. We see that each time there is a closed loop of the solid lines we get a factor N. Thus if we take the limit $N \to \infty$, $\lambda_0 \to 0$ keeping $\lambda_0 N$ fixed, only a subset of the diagrams like those in Fig. 29 survives.

If we denote the sum of these by $G^{(2)}(p)$, it satisfies the equation

$$G^{(2)}(p) = \frac{1}{p^2 + m_0^2} - \frac{\lambda_0 N}{2}\frac{1}{p^2 + m_0^2} G^{(2)}(p) \int G^{(2)}(k) \frac{d^d k}{(2\pi)^d}$$

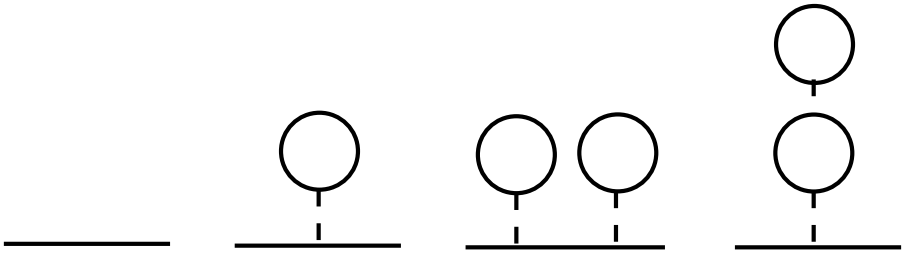

Figure 29: The first few surviving diagrams in the limit $N \to \infty$ with $\lambda_0 N$ fixed.

Equivalently

$$\Gamma^{(2)}(p) = p^2 + m_0^2 + \lambda_0 N \int \frac{1}{\Gamma^{(2)}(k)} \frac{d^d k}{(2\pi)^d}$$

Note that the last term is independent of p, so we can write

$$\Gamma^{(2)}(p) = p^2 + m^2$$

with

$$m^2 = m_0^2 + \frac{\lambda_0 N}{2} \int \frac{1}{k^2 + m^2} \frac{d^d k}{(2\pi)^d}$$

This is an equation which gives the renormalised mass m implicitly in terms of the bare mass. The renormalised mass vanishes when

$$m_0^2 = m_{0c}^2 = -\frac{\lambda_0 N}{2} \int \frac{1}{k^2} \frac{d^d k}{(2\pi)^d}$$

Note this is negative as we mentioned earlier, corresponding to the fact that $T_c < T_{MF}$. The integral is UV divergent for $d \geq 2$, but we expect this, and it can be made finite with a regulator $|k| < \Lambda$. However, it is also IR divergent for $d \leq 2$, indicating that, even in the UV cutoff theory, starting with any finite value of the bare mass, we can never in fact reach $m = 0$. This is an example of Coleman's theorem (also called the Mermin-Wagner theorem in statistical physics), which states that it is impossible to break a continuous symmetry for $d \leq 2$. As $d \to 2+$, $T_c \to 0$.

Introducing the notation $\delta T = m_0^2 - m_{0c}^2$ as before, we see that

$$\begin{aligned} m^2 &= \delta T + \frac{\lambda_0 N}{2} \int \frac{1}{k^2 + m^2} \frac{d^d k}{(2\pi)^d} - \frac{\lambda_0 N}{2} \int \frac{1}{k^2} \frac{d^d k}{(2\pi)^d} \\ &= \delta T - \frac{\lambda_0 N}{2} m^2 \int \frac{1}{k^2(k^2 + m^2)} \frac{d^d k}{(2\pi)^d} \end{aligned}$$

There are different types of behaviour to the solution depending on d:

- $d > 4$: in this case the integral must still be regulated. The integral on the rhs is then finite at $m = 0$, and we see that $m^2 \propto (\delta T)$. Thus the critical exponent $\nu = \frac{1}{2}$, independent of d. This is typical of the behaviour above the *upper critical dimension*, in this case 4.

- $2 < d < 4$: the integral on the rhs is finite, and, by dimensional analysis, goes like m^{d-4}. The leading behaviour of the solution is now found by balancing the two terms on the rhs, whence $m \propto (\delta T)^{1/(d-2)}$, and $\nu = 1/(d-2)$.
- for $d \leq 2$, as we said above, there is no solution.

In addition one can check that there are logarithmic corrections at $d = 4$. These general features, if not the precise values of ν, persist to finite N.

7 From Feynman diagrams to Cross-sections

In this section we assume that we are in 3+1-dimensional Minkowski space.

Earlier, we saw that in a free field theory

$$\langle 0|\mathbf{T}[\hat{\phi}(\mathbf{p}',t')\hat{\phi}(\mathbf{p},t)^\dagger]|0\rangle = (2\pi)^3\delta^3(\mathbf{p}'-\mathbf{p})\frac{e^{-i\sqrt{\mathbf{p}^2+m^2}(t'-t)}}{2\sqrt{\mathbf{p}^2+m^2}} \qquad (18)$$

is proportional to the amplitude $e^{-i\sqrt{\mathbf{p}^2+m^2}(t'-t)}$ for a particle of 3-momentum $\mathbf{p}$ at time t to propagate to one of momentum $\mathbf{p}'$ at time t'. That is $\hat{\phi}(\mathbf{p})^\dagger$ acting on the vacuum state creates a single particle state $|\mathbf{p}\rangle$ of momentum $\mathbf{p}$. More precisely,

$$\langle \mathbf{p}'|\hat{\phi}(\mathbf{p})^\dagger|0\rangle = \delta(\mathbf{p}-\mathbf{p}')\,.$$

The extra factor of $(2\sqrt{\mathbf{p}^2+m^2})^{-1}$ in (18) is accounted for by the fact that in relativistic quantum mechanics, it is standard to use a Lorentz invariant normalisation of the momentum eigenstates. That is, rather than the completeness relation

$$\int \frac{d^3\mathbf{p}}{(2\pi)^3}|\mathbf{p}\rangle\langle\mathbf{p}| = 1$$

we use

$$\int \frac{d^4p}{(2\pi)^3}\delta^+(p^2-m^2)|\mathbf{p}\rangle\langle\mathbf{p}| = \int \frac{d^3\mathbf{p}}{(2\pi)^3(2p_0)}|\mathbf{p}\rangle\langle\mathbf{p}| = 1$$

where the superscript + means that we take only the zero of the delta function with $p_0 > 0$.

Now consider scattering in an interacting field theory. In principle we should try to model what goes on at CERN and consider isolated wave packets prepared at $t \to -\infty$, allowed to interact around $t = 0$, and then see what state comes out as $t \to +\infty$. In practice this is too hard and instead we suppose that we have a large box and that as both $t \to \pm\infty$ the interaction is switched off. Initially, in the Schrödinger picture, we have a plane wave state $|\mathbf{p}_1, \mathbf{p}_2, \ldots\rangle^{(S)}$. This evolves with the full hamiltonian $\hat{H}^{(S)}$, and then, as $t \to +\infty$, we compute the transition amplitude

$$\lim_{t\to\infty} {}^{(S)}\langle \mathbf{p}'_1, \mathbf{p}'_2, \ldots | e^{-i\hat{H}^{(S)}t} | \mathbf{p}_1, \mathbf{p}_2, \ldots\rangle^{(S)} = {}^{(S)}\langle \mathbf{p}'_1, \mathbf{p}'_2, \ldots | \hat{S} | \mathbf{p}_1, \mathbf{p}_2, \ldots\rangle^{(S)}$$

that tells us the probability of scattering from one plane-wave state into another. $\hat{S}$ is called the S-matrix.

The way to do this is through the *interaction picture*. We write $\hat{H}^{(S)} = \hat{H}_0^{(S)} + \hat{H}_1^{(S)}$, where the first term is the free, or non-interacting hamiltonian, and the second part the interaction terms. Going from the Schrödinger to the interaction picture corresponds to a unitary transformation

$$\begin{aligned} |\Psi\rangle^{(I)} &= e^{i\hat{H}_0^{(S)}t} |\Psi\rangle^{(S)} \quad \text{on states} \\ \hat{\mathcal{O}}^{(I)} &= e^{i\hat{H}_0^{(S)}t} \hat{\mathcal{O}}^{(S)} e^{-i\hat{H}_0^{(S)}t} \quad \text{on operators} \end{aligned}$$

Then

$$\begin{aligned} i\frac{\partial}{\partial t}|\Psi\rangle^{(I)} &= i\left[i\hat{H}_0^{(S)} e^{i\hat{H}_0^{(S)}t} |\Psi\rangle^{(S)} + e^{i\hat{H}_0^{(S)}t} \frac{\partial}{\partial t} |\Psi\rangle^{(S)} \right] \\ &= -\hat{H}_0^{(S)} |\Psi\rangle^{(I)} + e^{i\hat{H}_0^{(S)}t} \left(\hat{H}_0^{(S)} + \hat{H}_1^{(S)} \right) e^{-i\hat{H}_0^{(S)}t} e^{i\hat{H}_0^{(S)}t} |\Psi\rangle^{(S)} \\ &= \hat{H}_1^{(I)} |\Psi\rangle^{(I)} \end{aligned}$$

We define the $\hat{U}$ operator by

$$|\Psi(t)\rangle^{(I)} = \hat{U}(t, t_0) |\Psi(t_0)\rangle^{(I)}$$

Note that $\hat{U}(t_0, t_0) = 1$ and $\hat{S}^{(I)} = \hat{U}(\infty, -\infty)$. From the above $\hat{U}$ obeys the equation

$$i(\partial/\partial t)\hat{U}(t, t_0) = \hat{H}_1^{(I)}(t)\hat{U}(t, t_0)$$

It is the dependence of $H_1^{(I)}$ on t which makes the exact solution of this equation difficult. However we can write it as

$$\hat{U}(t, t_0) = 1 - i\int_{t_0}^{t} dt' H_1^{(I)}(t')\hat{U}(t', t_0)$$

and find a solution by iteration:

$$\begin{aligned}
\hat{U}(t,t_0) &= 1 - i\int_{t_0}^{t} dt_1 \hat{H}_1^{(I)}(t_1)\left[1 - i\int_{t_0}^{t_1} dt_2 \hat{H}_1^{(I)}(t_2)\hat{U}(t_2,t_0)\right] \\
&= \sum_{n=0}^{\infty}(-i)^n \int_{t_0}^{t} dt_1 \int_{t_0}^{t_1} dt_2 \ldots \int_{t_0}^{t_{n-1}} dt_n \hat{H}_1^{(I)}(t_1)\hat{H}_1^{(I)}(t_2)\ldots\hat{H}_1^{(I)}(t_n) \\
&= \sum_{n=0}^{\infty}\frac{(-i)^n}{n!} \int_{t_0}^{t} dt_1 \int_{t_0}^{t} dt_2 \ldots \int_{t_0}^{t} dt_n \mathbf{T}\left[\hat{H}_1^{(I)}(t_1)\hat{H}_1^{(I)}(t_2)\ldots\hat{H}_1^{(I)}(t_n)\right] \\
&= \mathbf{T}\left[\exp\left(-i\int_{t_0}^{t} dt \hat{H}_1^{(I)}(t)\right)\right]
\end{aligned}$$

where the last line is just shorthand for the one above. Putting all this together, the transition amplitude is

$$\begin{aligned}
&{}^{(S)}\langle \mathbf{p}_1', \mathbf{p}_2', \ldots | \hat{S}^{(S)} | \mathbf{p}_1, \mathbf{p}_2, \ldots \rangle^{(S)} \\
&= {}^{(I)}\langle \mathbf{p}_1', \mathbf{p}_2', \ldots | \hat{S}^{(I)} | \mathbf{p}_1, \mathbf{p}_2, \ldots \rangle^{(I)} \\
&= {}^{(I)}\langle \mathbf{p}_1', \mathbf{p}_2', \ldots | \mathbf{T}\left[\exp\left(-i\int_{-\infty}^{\infty} dt \hat{H}_1^{(I)}(t)\right)\right] | \mathbf{p}_1, \mathbf{p}_2, \ldots \rangle^{(I)} \\
\propto\ &\langle 0 | \mathbf{T}\left[\left(\prod_{j'} \hat{\phi}(p_{j'}')^{(I)}\right)\exp\left(-i\int_{-\infty}^{\infty} dt \hat{H}_1^{(I)}(t)\right)\left(\prod_{j} \hat{\phi}^{\dagger}(p_j)^{(I)}\right)\right] | 0 \rangle
\end{aligned}$$

In the path integral formulation, if the operators on the left and right were in the Schrodinger picture, this would correspond to

$$\langle \prod_{j'} \phi(p_{j'}') \exp\left(i\int d^4x \mathcal{L}_1(x)\right) \prod_{j} \phi(p_j) \rangle_0$$

in the free theory, since $H_1(x) \to -\int d^3x \mathcal{L}_1(\mathbf{x})$. This is just the correlation function

$$G^{(N'+N)}(\{p_{j'}'\},\{p_j\}) = \langle \prod_{j'} \phi(p_{j'}') \prod_{j} \phi(p_j) \rangle$$

(times an overall energy-momentum conserving delta-function) in the *interacting theory.*

However there is a slight difference. If we switch off the interaction, that is set $\hat{H}_1 = 0$, then the states in the interaction picture do not evolve. That is $\hat{S} = 1$ or

$${}^{(S)}\langle \mathbf{p}_1', \mathbf{p}_2', \ldots | \hat{S}^{(S)} | \mathbf{p}_1, \mathbf{p}_2, \ldots \rangle^{(S)} = \delta(\{p_{j'}'\},\{p_j\})$$

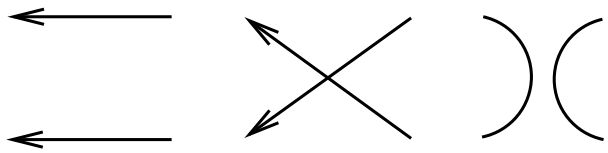

Figure 30: Disconnected diagrams contributing to the $2 \to 2$ S-matrix. The arrows indicate the flow of energy (p_0). The third diagram does not contribute.

However in, *e.g.*, the $2 \to 2$ case, the diagrams contributing to $G^{(2+2)}$ are the first two shown in Fig. 30 (why not the third diagram?), giving

$$\frac{1}{(p_1^2 - m^2)(p_2^2 - m^2)} \left(\delta_{p_1 p_1'} \delta_{p_2 p_2'} + \delta_{p_1 p_2'} \delta_{p_2 p_1'} \right)$$

Thus we should cancel off the denominators in the above, to be in the interaction picture.

When we turn on the interaction, we can write

$$\hat{S} = 1 + i\hat{\mathcal{T}}$$

The 1 in $\hat{S}$ now corresponds, for $2 \to 2$ scattering, to the sum of disconnected diagrams in $G^{(2)}(p_1)G^{(2)}(p_2)$, where $G^{(2)}(p)$ is the full renormalised 2-point function. Note that we are now saying that the physical particle states are created and destroyed by the *renormalised* field operator. By the above argument, we should however multiply this by

$$(p_1^2 - m^2)(p_2^2 - m^2)G^{(2)}(p_1)G^{(2)}(p_2)$$

and then take the limit $p_1^2, p_2^2 \to m^2$, since the external particles are on mass shell. However, since the field renormalisation is defined by the requirement that the residue of the pole in $G^{(2)}(p)$ at $p^2 = m^2$ is 1, we get just the sum of energy-momentum conserving delta-functions as before.

However the 1 term in $\hat{S}$ corresponds to no scattering, or to looking exactly in the forward direction, down the beamline at CERN. To see scattering we need to look at the other term. The first diagram contributing to the matrix element of $\hat{\mathcal{T}}$ in ϕ^4 theory is that in Fig. 31, with all possible vacuum corrections on the external legs. Once again, however, the corrections are cancelled on mass shell by renormalisation, and the remaining external propagators are cancelled by the use of the interaction picture. That is,

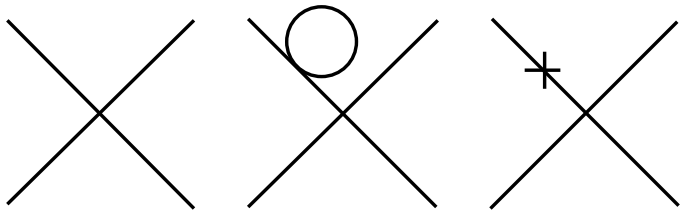

Figure 31: Diagrams contributing to the $2 \to 2$ $\mathcal{T}$-matrix to $O(\lambda)$. The loop corrections on the extenral legs are exactly cancelled by mass and field renormalisation.

the amplitude for the 2 particles to propagate up to the time where they first interact is unity in this picture. This is an example of the main important result:

The matrix elements of $i\hat{\mathcal{T}}$ are given by the on-mass-shell values of the truncated renormalised connected N-point functions

$$i\langle \mathbf{p}'_1, \mathbf{p}'_2, \ldots | \hat{\mathcal{T}} | \mathbf{p}_1, \mathbf{p}_2, \ldots \rangle = \left. \frac{G^{(N'+N)}(\{p'_{j'}\}, \{p_j\})_c}{\Pi_{j'} G^{(2)}(p'_{j'}) \, \Pi_j G^{(2)}(p_j)} \right|_{{p'_{j'}}^2 = p_j^2 = m^2} \times \delta^4(\sum_{j'} p'_{j'} - \sum_j p_j)$$

We have not attempted to show this in general (the full argument requires the use of the so-called LSZ formalism and is cumbersome.) For example, in $\lambda\phi^4$ theory, we have simply

$$\langle \mathbf{p}'_1, \mathbf{p}'_2 | \hat{\mathcal{T}} | \mathbf{p}_1, \mathbf{p}_2 \rangle = \lambda \, (2\pi)^4 \delta^4(p'_1 + p'_2 - p_1 - p_2) + O(\lambda^2)$$

The transition probability is the square of this, which involves

$$\left[\delta^4(p'_1 + p'_2 - p_1 - p_2)\right]^2 = \delta^4(0)\delta^4(p'_1 + p'_2 - p_1 - p_2)$$

If we are careful, we should identify the divergent quantity $(2\pi)^4\delta^4(0)$ as VT where V is the volume of the box, and T the total time for which the interaction is switched on. We divide this by T to get a transition *rate*. If we are in a box, the allowed 3-momenta are quantised in units of $(2\pi)^3/V$. In order to go from our relativistic normalisation of the states

$$\langle \mathbf{p}' | \mathbf{p} \rangle = 2p^0 \delta^3(\mathbf{p} - \mathbf{p}')$$

to discrete counting where

$$\langle \mathbf{p}' | \mathbf{p} \rangle = \delta_{\mathbf{p},\mathbf{p}'}$$

that is

$$\int \frac{d^3p}{(2\pi)^3} \to \frac{1}{2p_0 V} \sum_{\mathbf{p}},$$

we need to divide the transition rate by $\prod_{j'}(2p'^0_j V)\prod_j(2p^0_j V)$.

According to Fermi's golden rule, we now multiply by the density of final states

$$\frac{V d^3p'_1}{(2\pi)^3}\frac{V d^3p'_2}{(2\pi)^3}$$

and normalise against the total incident flux v_{12}/V, where v_{12} is the relative velocity, to get the elastic differential cross-section

$$d\sigma = \frac{V}{v_{12}} \frac{V(2\pi)^4\delta^4(p'_1+p'_2-p_1-p_2)}{\prod_{j'}(2p'^0_j V)\prod_j(2p^0_j V)} \frac{V d^3p'_1}{(2\pi)^3}\frac{V d^3p'_2}{(2\pi)^3}\ \left(\lambda^2+\cdots\right)$$

Note that all factors of V cancel.

In the centre of mass frame we can work out $p_1, \ldots, p'_2$ in terms of the total centre of mass energy $E = 2\sqrt{\mathbf{p}_1^2 + m^2}$ and the scattering angle θ. After some algebra we find that

$$\delta^4(p'_1+p'_2-p_1-p_2)d^3p'_1 d^3p'_2 \propto \sin\theta d\theta d\phi = d\Omega$$

and the differential cross-section is

$$\left.\frac{d\sigma}{d\Omega}\right|_{\rm CM} = \frac{|T|^2}{64\pi^2 s}$$

where $s = (p_1+p_2)^2 = E^2$. This formula is true in general for the elastic scattering of scalar particles of equal mass. In our case $T = \lambda + O(\lambda)^2$. Note that to lowest order this is independent of θ (there is only S-wave scattering) but to higher orders this is no longer the case.

7.1 The S-matrix: analyticity and unitarity

We have shown that it useful to write $\hat{S} = 1 + i\hat{\mathcal{T}}$. However $\hat{S}$ is unitary, $\hat{S}\hat{S}^\dagger = \hat{S}^\dagger\hat{S} = 1$, so

$$\hat{\mathcal{T}}^\dagger\hat{\mathcal{T}} = i(\hat{\mathcal{T}}^\dagger - \hat{\mathcal{T}})$$

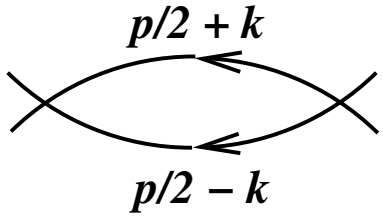

Figure 32: One loop contribution to $\hat{\mathcal{T}}$. To get the imaginary part, put the internal lines on mass shell.

Thus, although we saw, in our example, that to lowest order $\hat{\mathcal{T}} \sim \lambda$ is real, this shows that to higher orders it has an imaginary part. In fact if we insert a complete set of multiparticle states $|\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3, \ldots\rangle$, the diagonal matrix element of left hand side between 2-particle states is

$$\sum_{\{\mathbf{k}_j\}} \langle \mathbf{p}_1, \mathbf{p}_2 | \hat{\mathcal{T}}^\dagger | \{\mathbf{k}_j\}\rangle\langle\{\mathbf{k}_j\}|\hat{\mathcal{T}}|\mathbf{p}_1, \mathbf{p}_2\rangle$$

The summand here is the square of the transition amplitude, that is the cross-section for $p_1 + p_2 \to \sum_j k_j$. Hence the lhs is proportional to the total cross-section, $\sigma_{\rm tot}$, for 2 particles $\to$ anything. The rhs, on the other hand is the imaginary part of the forward (that is $(p_1', p_2') = (p_1, p_2)$) scattering amplitude. This is the *optical theorem*.

It is interesting to see how this works in $\lambda\phi^4$ theory. The $O(\lambda^2)$ contribution to the matrix element of $\hat{\mathcal{T}}$ is given by Fig. 32. In the CM frame we have

$$M \equiv \langle \mathbf{p}_1', \mathbf{p}_2' | i\hat{\mathcal{T}} | \mathbf{p}_1, \mathbf{p}_2\rangle = i\lambda + \frac{(i\lambda)^2}{2}\int \frac{dk_0 d^3\mathbf{k}}{(2\pi)^4} \frac{i^2}{((k_0 - \frac{1}{2}E)^2 - \mathbf{k}^2 - m^2 + i\epsilon)((k_0 + \frac{1}{2}E)^2 - \mathbf{k}^2 - m^2 + i\epsilon)}$$

where $E = p_1^0 + p_2^0$ is the total CM energy.

Now think about doing the k_0 integral by contour integration. The poles are at $k_0 = \frac{1}{2}E \pm \sqrt{\mathbf{k}^2 + m^2} \mp i\epsilon$ and $k_0 = -\frac{1}{2}E \pm \sqrt{\mathbf{k}^2 + m^2} \mp i\epsilon$. Their positions are shown in Fig. 33 for $E < 2m$. In that case we can complete the contour (in either half-plane): we pick up a factor i from Cauchy's theorem, so the overall contribution to M is still real. It is also analytic in E, since we can always move the contour away from the poles, and the integral is uniformly convergent.

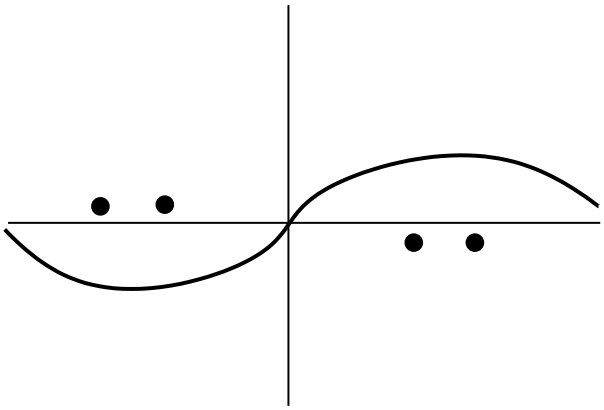

Figure 33: Position of the poles in the complex k_0-plane for $E < 2m$. The integration contour is indicated.

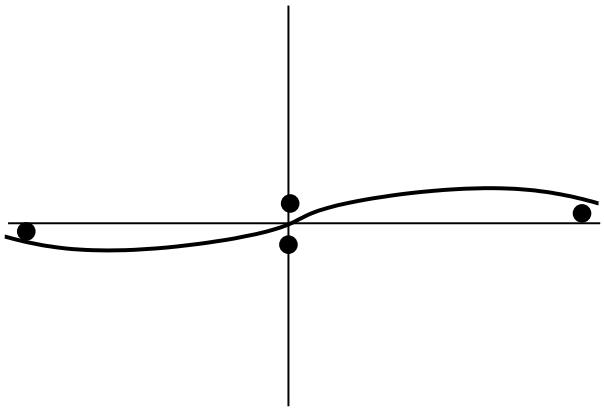

Figure 34: Position of the poles in the complex k_0-plane for $\mathrm{Im}\, E > 0$ and $\mathrm{Re}\, E$ continued $> 2m$. The two central poles pinch the integration contour as $\mathrm{Im}\, E$ and $\epsilon \to 0$.

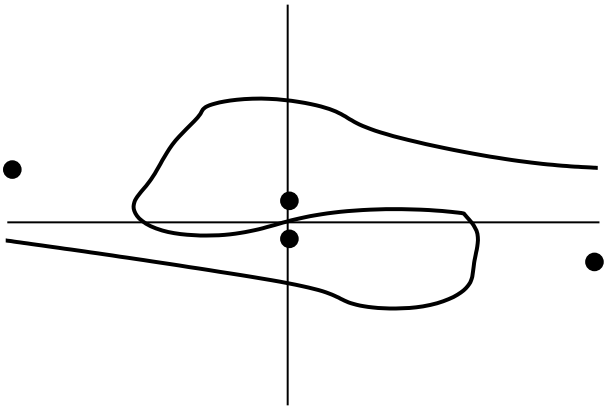

Figure 35: Position of the poles in the complex k_0-plane for $\mathrm{Im}\, E < 0$ and $\mathrm{Re}\, E$ continued $> 2m$. The two central poles now pinch the contour in the opposite sense.

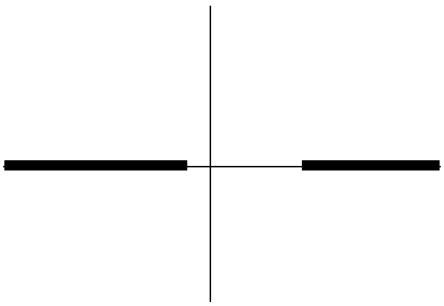

Figure 36: Analytic properties of the $\mathcal{T}$-matrix in the complex s-plane. There are branch cuts starting at $s = 4m^2$ and $s = -t$. In between, the function is real on the real axis.

However if we try to increase E past $2m$ the poles (almost, apart from the $i\epsilon$s) collide. If we first give E an imaginary part $E \to E + i\delta$ and then increase the real part through $2m$, then, for a given $\mathbf{k}$, the contour gets pinched as $\delta, \epsilon \to 0$. For $\delta \to 0+$ it gets pinched as in Fig. 34, for $\delta \to 0-$ as in Fig. 35. This tells us that M is non-analytic as $\delta \to 0$, and in fact gets an *imaginary part*. If we take the path with $\delta > 0$ and then let $\delta \to 0$ we end up with an imaginary part $+\mathrm{Im}\, M$; if we take the path with $\delta < 0$ we get $-\mathrm{Im}\, M$, so there is a discontinuity $2i\mathrm{Im}\, M$. The magnitude of this is found by the values of the integrand where the poles collide, namely it is given by

$$\tfrac{1}{2}\lambda^2 \int \frac{dk_0 d^3\mathbf{k}}{(2\pi)^4} \delta^+((k_0 + \tfrac{1}{2}E)^2 - \mathbf{k}^2 - m^2)\, \delta^+((k_0 + \tfrac{1}{2}E)^2 - \mathbf{k}^2 - m^2)$$

This corresponds to inserting a complete set of intermediate states (only 2-particle states are allowed to this order) and so gives $\hat{\mathcal{T}}^\dagger \hat{\mathcal{T}}$, consistent with the optical theorem.

The analyticity properties of the matrix elements M of $\hat{\mathcal{T}}$ we saw in this example are in fact quite general and very powerful. It is usual to express M in terms of the relativistic invariant $s \equiv (p_1 + p_2)^2 = E_{\mathrm{CM}}^2$. Then $M(s)$ is an analytic function of s, real on the real axis for $0 < s < 4m^2$, but it has a branch cut beginning at $s = 4m^2$ (see Fig. 36). For $s > 4m^2$ the physical value of M is given by its value just above the cut. The point $s = 4m^2$ is called the 2-particle threshold. Of course for physical on-shell incident particles we must have $E_{\mathrm{CM}} \geq 2m$, so $s \geq 4m^2$. This corresponds to the value of s at which $2 \to 2$ scattering is possible. However there

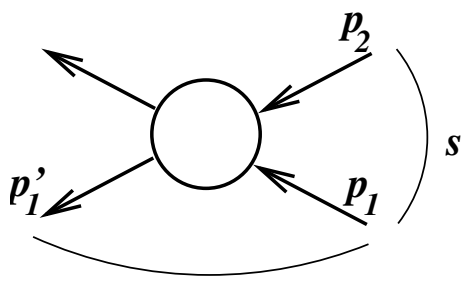

Figure 37: The s and t-channels for $2 \to 2$ scattering.

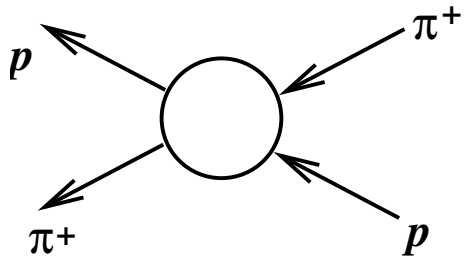

Figure 38: The $\mathcal{T}$-matrix for $\pi^+ p \to \pi^+ p$ scattering in the s-channel, when analytically continued to the u-channel, describes $p\bar{p} \to \pi^+\pi^-$.

are higher thresholds at $s = 9m^2$, etc., where $2 \to 3$ particle production becomes possible. These add even more branch cuts on top of the first one. However, we can say more than this. In general M is an analytic function of the relativistic invariants $s = (p_1+p_2)^2$, $t = (p_1-p_1')^2$, and $u = (p_1-p_2')^2$ (see Fig. 37). It is easy to check that

$$s + t + u = 4m^2$$

and, just as there is a branch cut starting at $s = 4m^2$ there is also one starting at $u = 4m^2$, that is $s = -t$. So for fixed $t > 0$ the analytic structure of $M(s)$ is as shown in Fig. 36. The physical value in the u-channel, that is for scattering $1+2' \to 1'+2$, is the one *below* the left-hand branch cut in the s-plane.

This is important for the scattering of real particles. If, for example, the s-channel corresponds to elastic pion-proton scattering $\pi^+ p \to p\pi^+$, then the u-channel corresponds to proton-antiproton annihilation into a $\pi^+\pi^-$ pair. See Fig. 38. The fact that the amplitude for one is the analytic continuation of the other means that in principle we can predict the cross-

section for the second process if we know the amplitude for the first.

8 Path integrals for fermions

We have seen that correlation functions computed with the path integral correspond to vacuum expectation values of time-ordered products of quantum field operators, for example,

$$\begin{aligned}\langle\phi(x_1)\phi(x_2)\phi(x_3)\phi(x_4)\rangle &= Z^{-1}\int[d\phi]\phi(x_1)\phi(x_2)\phi(x_3)\phi(x_4)\,e^{iS[\phi]} \\ &= \langle 0|\mathbf{T}[\hat{\phi}(x_1)\hat{\phi}(x_2)\hat{\phi}(x_3)\hat{\phi}(x_4)]|0\rangle\end{aligned}$$

When $t_1 = t_2 > t_3 = t_4$ we interpret this as creating particles at x_3 and x_4 at time t_3 and destroying them at t_1. However in the path integral the $\phi(x_j)$ are commuting c-numbers, so that

$$\langle\phi(x_1)\phi(x_2)\phi(x_3)\phi(x_4)\rangle = \langle\phi(x_2)\phi(x_1)\phi(x_4)\phi(x_3)\rangle$$

This tells us that the initial and final quantum states are symmetric under the exchange of 3 and 4: the particles we are describing must be *bosons*.

How, then, are we to describe fermions using the path integral? The answer is to define a new kind of integral over *anti-commuting* fields, also called a Grassmann integral. The idea is that the only mathematical properties of ordinary integration that we actually need to use in manipulating ordinary path integrals are the fact that it is a map from functions to the complex numbers which is linear:

$$\int(\alpha f(x) + \beta g(x))dx = \alpha\int f(x)dx + \beta\int g(x)dx$$

and that it satisfies the conditions that we can shift variables and rescale variables in integrals from $-\infty$ to ∞:

$$\int_{-\infty}^{\infty} dx\, f(x+c) = \int_{-\infty}^{\infty} dx\, f(x) \tag{19}$$

$$\int_{-\infty}^{\infty} dx\, f(\alpha x) = \alpha^{-1}\int_{-\infty}^{\infty} dx\, f(x) \tag{20}$$

Let first consider a finite set of anti-commuting numbers $(\theta_1, \ldots, \theta_n)$, satisfying

$$\theta_j \theta_k = -\theta_k \theta_j$$

Note that this immediately implies that $\theta_j^2 = 0$, and therefore the Taylor expansion of any function $f(\theta_1, \ldots, \theta_n)$ *terminates* after a finite number of terms, for example

$$\begin{aligned} f(\theta) &= f_0 + f_1 \theta \\ f(\theta_1, \theta_2) &= f_{00} + f_{10}\theta_1 + f_{01}\theta_2 + f_{11}\theta_1\theta_2 \end{aligned}$$

Now let us try to define the 'integral' of $f(\theta)$ as a linear map from functions to the complex numbers. It should satisfy the shifting property (19)

$$\int d\theta\, f(\theta + \theta') = \int d\theta\, f(\theta)$$

that is

$$\int d\theta (f_0 + f_1\theta + f_1\theta') = \int d\theta (f_0 + f_1\theta)$$

This means that

$$\int d\theta\, 1 = 0$$

We are in fact free to choose the value of $\int d\theta\, \theta$ and usually set

$$\int d\theta\, \theta = 1$$

Note however that

$$\int d\theta\, f(\alpha\theta) = \int d\theta\, f_1 \alpha\theta = \alpha \int d\theta\, f(\theta)$$

as compared with (20) for the case of ordinary integration.

In path integrals we need to understand how to go gaussian integration. Consider therefore

$$\int d\theta_1 \ldots d\theta_n\, e^{-\frac{1}{2}\sum_{j,k=1}^n \theta_j M_{jk} \theta_k}$$

(Note that the matrix M has to be antisymmetric for this to make sense.) Take the case $n = 4$. Expanding out the exponential we have

$$\int d\theta_1 \ldots d\theta_4 \left[1 - \tfrac{1}{2} \sum_{jk} \theta_j M_{jk} \theta_k + \tfrac{1}{8} (\sum_{jk} \theta_j M_{jk} \theta_k)(\sum_{j'k'} \theta_{j'} M_{j'k'} \theta_{k'}) + O((\sum_{jk} \theta_j M_{jk} \theta_k)^3) \right]$$

The first two terms don't have enough factors of $\theta_1 \dots \theta_4$, and they integrate to zero. The 4th and higher terms all have at least one θ_j raised to a power ≥ 2, and so vanish. The remaining contribution is

$$\begin{aligned}
&\int d\theta_1 \dots d\theta_4 \; [\theta_1\theta_2\theta_3\theta_4 M_{12}M_{34} + \theta_1\theta_3\theta_2\theta_4 M_{13}M_{24} + \cdots] \\
&\propto \int d\theta_1 \dots d\theta_4 \, \theta_1\theta_2\theta_3\theta_4 \, [M_{12}M_{34} - M_{13}M_{24} + \cdots] \\
&= M_{12}M_{34} - M_{13}M_{24} + M_{14}M_{23} \\
&= \sqrt{\det M}
\end{aligned}$$

This is to be compared with the ordinary gaussian integral

$$\int dx_1 \dots dx_n \, e^{-\frac{1}{2}\sum_{j,k=1}^n x_j M_{jk} x_k} \propto (\det M)^{-1/2}$$

The main point, however, is that if we add sources:

$$Z[J] \equiv \int d\theta_1 \dots d\theta_n \, e^{-\frac{1}{2}\sum_{j,k=1}^n \theta_j M_{jk} \theta_k + \sum_j J_j \theta_j}$$

(note that the J's anticommute as well!) we can complete the square as before to find that

$$Z[J] = \sqrt{\det M} \, e^{\frac{1}{2}\sum_{jk} J_j (M^{-1})_{jk} J_k}$$

Taking derivatives wrt the sources we then find that

$$\begin{aligned}
\langle \theta_1\theta_2\theta_3\theta_4 \rangle &= (M^{-1})_{12}(M^{-1})_{34} - (M^{-1})_{13}(M^{-1})_{24} + (M^{-1})_{14}(M^{-1})_{23} \\
&= \langle \theta_1\theta_2 \rangle \langle \theta_3\theta_4 \rangle - \langle \theta_1\theta_3 \rangle \langle \theta_2\theta_4 \rangle + \langle \theta_1\theta_4 \rangle \langle \theta_2\theta_3 \rangle
\end{aligned}$$

This is the fermionic version of Wick's theorem: note the minus sign!

8.0.1 Relativistic lagrangians

By analogy with the bosonic case, our first guess at writing down a Lorentz-invariant action for a field theory describing a single species of fermion might be

$$\mathcal{L} \propto \psi(-\partial^2 + m^2)\psi$$

where $\psi(x)$ is an anticommuting field. But this is zero because the operator $-\partial^2 + m^2$ is symmetric. If it has to be antisymmetric then we need first

derivatives, like $\psi\partial\psi$. But this is not Lorentz invariant. One way to solve this is by adding more components to the field. For example in $d = 4$ we can give it 2 components (ψ_1, ψ_2) and we can then show that the Weyl lagrangian

$$\mathcal{L}_W = \sum_{j,k=1}^{2} \psi_j(\sigma_\mu \partial^\mu)\psi_k$$

where $\sigma_\mu = (1, \sigma)$ and σ are the Pauli matrices, is indeed Lorentz invariant, as long as (ψ_1, ψ_2) transforms like a spinor. This describes massless spin-$\frac{1}{2}$ particles (like massless neutrinos.)

Note that it is still hard to add a mass term. The easiest way is to double the number of components again:

$$\begin{aligned} \mathcal{L}_D &= \psi_L^T(\sigma \cdot \partial)\psi_L + \psi_R^T(\sigma \cdot \partial)\psi_R + im\psi_L^T\psi_R \\ &= \overline{\Psi}(\gamma \cdot \partial - m)\Psi \end{aligned}$$

where $\Psi = (\psi_L, \psi_R)$, and $\overline{\Psi} = \Psi^T\gamma_0$. This gives free massive spin-$\frac{1}{2}$ particles, and their antiparticles, satisfying the Dirac equation. Note that the propagator for such a particle in field theory is $(i\gamma \cdot p + m)^{-1}$, which is a 4×4 *matrix*.

8.0.2 An application in condensed matter physics

Grassmann integrals are also very useful in condensed matter physics, not only in describing non-relativistic electrons.

Suppose that we have a single non-relativistic particle in a potential, whose wave function satisfies the time-independent Schrodinger equation

$$\hat{H}\phi = (-\partial^2 + V(r))\phi(r) = E\phi(r)$$

We are interested in the Green's function

$$G(r_1, r_2; E) = \left[\hat{H} - E\right]^{-1}_{r_1, r_2}$$

which we can write as a (bosonic) path integral

$$G(r_1, r_2; E) = Z^{-1} \int [d\phi]\phi(r_1)\phi(r_2)^* e^{-\frac{1}{2}\int \phi^*(-\partial^2 + V(r) - E)\phi d^d r}$$

where

$$Z = \int [d\phi][d\phi^*] e^{-\frac{1}{2}\int \phi^*(-\partial^2 + V(r) - E)\phi d^d r} = (\det(-\partial^2 + V(r) - E))^{-1}$$

(Note that we get det^{-1} rather than $det^{-1/2}$ because a complex field has two components. In systems containing impurities, V is random, and in many cases it is sufficient to calculate average properties of G (or $|G|^2$) over some ensemble of random potentials. This is difficult because V occurs in both the numerator and denominator. A way around this is to observe that if we consider the corresponding Grassmann integral

$$\int [d\theta][d\theta^*] e^{-\frac{1}{2}\int \theta^*(-\partial^2 + V(r) - E)\theta d^d r}$$

we get $Z = (\det(-\partial^2 + V(r) - E))^{+1}$. Hence we can write

$$\begin{aligned} G &= \int [d\phi][d\phi^*]\phi(r_1)\phi(r_2)^* e^{-\frac{1}{2}\int \phi^*(-\partial^2 + V(r) - E)\phi d^d r} \cdot \int [d\theta][d\theta^*] e^{-\frac{1}{2}\int \theta^*(-\partial^2 + V(r) - E)\theta d^d r} \\ &= \int [d\phi][d\phi^*][d\theta][d\theta^*]\phi(r_1)\phi(r_2)^* e^{-\frac{1}{2}\int \left(\phi^*(-\partial^2 + V(r) - E)\phi + \theta^*(-\partial^2 + V(r) - E)\theta\right) d^d r} \end{aligned}$$

In this form it is easy to perform an average over V. For example, if it has a gaussian distribution we get for each value of r,

$$\int e^{-\left(V^2/2\sigma + V(\phi^*\phi + \theta^*\theta)\right)} dV \propto e^{(\sigma/2)\left(\phi^*\phi + \theta^*\theta\right)^2}$$

The action for this field theory is invariant under global 'rotations' which take ϕ into ψ:

$$\begin{aligned} \phi(r) &\rightarrow \phi(r) + \epsilon\theta^*(r) \\ \theta(r) &\rightarrow \theta(r) - \epsilon\phi^*(r) \end{aligned}$$

where ϵ^* is a Grassmann number. This is an example of a *supersymmetry*– non-relativistic particles in a random potential are described by a supersymmetric ϕ^4 field theory!