
Quantum Field Theory III - Perturbation Theory and Feynman Diagrams

Nic Ford

1 Introduction

This article is part of a series on physics for mathematicians, and the third in a sub-series on quantum field theory. It builds directly on the material in the first and second quantum field theory articles, and I'll be assuming familiarity with their content in what follows.

In the last article, we were concerned with scattering processes, that is, processes in which some number of particles collide with each other and produce some number of other particles. Specifically, as we'll briefly review in the next section, we arrived at a description that depended on vacuum expectation values of time-ordered products of field operators, the quantities we called $G_A^{(n)}(x_1, \dots, x_n)$. This was certainly progress, but we left an obvious question unanswered: how do we actually compute $G^{(n)}$?

Our goal this time is to start answering that question. It turns out that it's not possible to compute $G^{(n)}$ exactly in any physically interesting situations. Luckily, there is an approximation procedure that, in many cases, is astonishingly effective, and is responsible for many of the incredibly precise physical predictions that have arisen from this framework. In this article we'll describe how that procedure works, and in particular how it leads to the famous *Feynman diagrams* as a central computational tool.

In the previous quantum field theory articles in this series, I've made reference to the existence of some mathematical problems underlying the theory. While we haven't always gone through the effort of resolving those problems, most of what we've said up until now can, at least in some sense, be made into an actual theorem that has an actual proof. Unfortunately, in the discussion we're about to have, the mathematical problems are going to become much more severe. Many of the mathematical objects we're about to discuss lack a coherent formal definition, and as a result the "naive" version of the computation we're about to lay out will result in answers that even physicists agree are nonsense. In the final section, we'll discuss these mathematical problems and set the stage for our strategy for resolving them. This strategy will be the main topic of the next article in this series.

Among the books I've read while putting together this series, the discussions I found most helpful for this particular article were the ones in Peskin and Schroeder's *An Introduction to Quantum Field Theory* and Folland's *Quantum Field Theory: A Tourist Guide for Mathematicians*. I'm grateful to Jordan Watkins for many helpful comments on an earlier version of this article.

2 Recap

Let's start by briefly recalling where we landed at the end of the last article, both as a refresher and to set some notation in place. That article was all about *scattering amplitudes* in quantum field theory, which describe experiments in which some number of particles arrive from far away, collide, and interact in some fashion, resulting in some number of new particles shooting out at the end. To describe these processes, we needed to construct *in states* and *out states* for the collections of incoming and outgoing particles; the scattering amplitude could then be expressed as the overlap between an in state and an out state.

This required the use of an *interpolating field* for the one-particle states, that is, an operator-valued distribution A with the property that

$$\langle p|A(x)|\Omega\rangle \neq 0,$$

where $|\Omega\rangle$ is the vacuum state and $|p\rangle$ is the one-particle state in question. We were able to show that, for any interpolating field, we'll always have

$$\langle p|A(x)|\Omega\rangle = Z^{1/2}e^{ip \cdot x},$$

where

$$Z^{1/2} = \langle p|A(0)|\Omega\rangle,$$

and that $Z^{1/2}$ is actually independent of the momentum p of the particle.

Given such an interpolating field, we defined the *time-ordered n -point function*

$$G_A^{(n)}(x_1, \dots, x_n) = \langle \Omega|T[A(x_1) \cdots A(x_n)]|\Omega\rangle,$$

where the *time-ordering symbol* T means that the factors appearing in brackets should be arranged in decreasing order in their time coordinates. We then defined the Fourier transform of $G^{(n)}$, which we called the *momentum-space time-ordered n -point function*, written

$$\tilde{G}_A^{(n)}(p_1, \dots, p_n) = \int d^4x_1 \cdots d^4x_n e^{-i \sum_{i=1}^n p_i \cdot x_i} G_A^{(n)}(x_1, \dots, x_n).$$

The culmination of the article was the proof of the *LSZ formula* for a scattering amplitude, which looked like

$$\langle p_{n+1}, \dots, p_{n+m} | p_1, \dots, p_n \rangle^{\text{out}} = \frac{(-i)^{n+m}}{Z^{(n+m)/2}} \prod_{k=1}^{n+m} (p_k^2 - \mu^2) \tilde{G}_A^{(n+m)}(p_1, \dots, p_n, -p_{n+1}, \dots, -p_{n+m}),$$

where μ is the mass of the particle. This formula is valid as long as all of the momenta of the incoming and outgoing particles are distinct.

In particular, we concluded from this that $\tilde{G}^{(n+m)}$ has a simple pole at 0 in each of the variables $p_k^2 - \mu^2$, and that our scattering amplitude can be thought of as the residue at that pole. Our job is now to explicitly compute the right side of this equation.

3 A Toy Lagrangian and Hamiltonian

In order to keep things concrete, we'll focus on one particular quantum field theory. Most of the essential features of the general case will appear in this example, and I think it's easier to digest the ones that don't after you've seen how the whole procedure works in a simple setting.

In the first article, we considered in some detail the quantization of a free real massive scalar field, which arose from the Lagrangian density

$$\mathcal{L}_0 = \frac{1}{2}((\partial\phi)^2 - m^2\phi^2).$$

(In the discussion that follows, we'll use a subscript 0 to indicate quantities that belong to a free field theory.) The theory we'll be working with arises from modifying this one by adding a term which is proportional to ϕ^4 , resulting in the Lagrangian density

$$\mathcal{L} = \frac{1}{2}((\partial\phi)^2 - m^2\phi^2) - \frac{\lambda}{4!}\phi^4.$$

(The factor of 4! is conventional; it's there to simplify part of the forthcoming Feynman diagram computation.)

This is called “ ϕ^4 theory,” and it's a standard first example in textbooks. Even though it doesn't describe any familiar physical processes, it strikes a nice balance between being simple enough that the computations aren't too overwhelming and being complicated enough that many of the interesting features of more physically realistic theories still show up.

We were able to explicitly produce a basis of eigenstates of the free Hamiltonian, essentially “solving” the theory, and as we are about to see this can be turned into a closed formula for the free-theory $\tilde{G}^{(n)}$'s without too much extra effort. This relied on the fact that the equations of motion were linear, a feature we no longer have, and in fact in essentially any interesting interacting quantum field theory there is no hope of finding an exact solution in the same way.

Instead, our strategy will have to be a bit more modest. We'll suppose that λ is small, and therefore regard \mathcal{L} as a small “perturbation” of the free Lagrangian \mathcal{L}_0 . Rather than try to compute $\tilde{G}^{(n)}$'s directly, our goal will instead be to compare them to their analogues for the free theory, producing a power series in λ which (we hope) will give us a decent approximation to the quantities we're actually interested in. Physicists refer to this procedure as **perturbation theory**.

Just as we did for the free theory, our first step is to produce a Hamiltonian. If you apply the same procedure we used for the free theory, you'll find that the Hamiltonian takes the form

$$H = \int d^3\mathbf{x} \left[\frac{1}{2} \left(\pi(\mathbf{x})^2 + |\nabla\phi(\mathbf{x})|^2 + m^2\phi(\mathbf{x})^2 \right) + \frac{\lambda}{4!}\phi(\mathbf{x})^4 \right].$$

It will be convenient to write

$$\mathcal{H}_I(\mathbf{x}) = \frac{\lambda}{4!}\phi(\mathbf{x})^4$$

and

$$H_I = \int d^3\mathbf{x} \mathcal{H}_I(\mathbf{x}),$$

so that

$$H = H_0 + H_I,$$

where H_0 is the **free Hamiltonian** that we considered in the first article. It's also common to refer to H_0 as the **free part** of H , and we'll call H_I the **interaction part** of H .

Recall that, in the free theory, $\phi(\mathbf{x})$ and $\pi(\mathbf{x})$ are not actually honest operators on the Hilbert space but rather *operator-valued distributions*. We therefore don't expect a product of two such “operators” to be meaningful, even in the distributional sense. You might recall that this caused a small problem even for the free Hamiltonian: we had to “subtract an infinite constant” from

H_0 in order to produce a *bona fide* operator on our Fock space. Shouldn't the presence of a term proportional to $\phi(\mathbf{x})^4$ make us even more worried?

The short answer is yes, it definitely should. In fact, if your goal is to construct a Hilbert space with operators on it corresponding to $\phi(\mathbf{x})$, $\pi(\mathbf{x})$, and H , this issue turns out to be fatal: as far as I know, the problem of realizing all of the computations we're about to perform using actual operators on an actual Hilbert space is a decades-old open problem that's not especially close to being resolved.

We'll have more to say about the mathematical issues in the final section of this article, but until then I invite you to think of the following story more as an analysis of the consequences of a series of conjectures. We are, sadly, not describing actual computations in an actual quantum-mechanical system; instead, it is more like we're describing how those computations *would* go if a system with the desired properties existed. Our strategy will therefore be to press on and manipulate expressions like " $\phi(\mathbf{x})^4$ " as though they referred to actual operators and see what we can learn. It's still possible to extract a lot of precise information in this way, and one of the remarkable features of quantum field theory is that, even in the absence of a mathematical foundation as firm as you might want, this information can still be compared with experiment.

With that in mind, let's get started.

4 The Perturbation Series

In this section, we're going to need to talk about states and operators that are associated with both the free theory and the interacting theory, so it will be helpful to have notation in place that reflects this. We've already introduced the symbols H_0 and H for the free part of the Hamiltonian and the full Hamiltonian, and we'll continue the practice introduced in the previous article of writing $|0\rangle$ for the vacuum state of the *free* theory and $|\Omega\rangle$ for the vacuum state of the *interacting* theory. (This distinction was introduced in a somewhat late revision to the previous article, so if you read it and this distinction between the two vacuum states is unfamiliar, that's why.)

That is, $|0\rangle$ the eigenstate of H_0 with the lowest eigenvalue, and similarly for $|\Omega\rangle$ and H . As a reminder, the claim that H has a unique lowest-energy state is, just like the existence of one-particle states in the interacting theory, a *hypothesis* of this setup — in both cases we're not claiming to have a proof that this property survives the change from H_0 to H , just that it is reasonable to hope it might if the perturbation is small enough.

4.1 Time Evolution Operators

In the free theory, the field operator $\phi(x)$ serves as an interpolating field for all of the one-particle states. Because we're thinking of our interacting theory as a small perturbation of the free theory, it therefore makes sense to hope that this might also be true there, so let's proceed under this assumption and try to compute scattering amplitudes for particles for which $\phi(x)$ is an interpolating field. To do this, we'll need to compute quantities of the form

$$\langle\Omega|T[\phi(x_1)\cdots\phi(x_n)]|\Omega\rangle.$$

(Actually, we will ultimately want the Fourier transform of this, but one step at a time!) Because we understand the free theory much better, our goal will be to relate this quantity to similar quantities computed in the free theory.

There's one small issue we won't be able to completely resolve right now. In order to satisfy our definition of "interpolating field" we need $\langle \Omega | \phi(x) | \Omega \rangle = 0$. When we discussed this condition in the last article, we said that if we don't have this for some candidate interpolating field we can fix it by subtracting a constant from the field operator, but that's actually not necessary for the example we're working with now — for ϕ^4 theory, there is a decent argument that our field operator satisfies this condition already. (Without a rigorous mathematical model, of course, "decent arguments" are the best we can get!) The reason is related to the fact that replacing $\phi(x)$ with $-\phi(x)$ is a symmetry of the Hamiltonian, but making this precise will have to wait until we've discussed how discrete symmetries like this are implemented in quantum field theory, which I hope to come to in a future article in this series. For now we'll just proceed under the assumption that this has been handled.

In the expressions for H_0 and H above, we are working in the "Schrödinger picture," where states depend on time and operators don't, and so ϕ and π are functions of space, but not time. If we switch to the "Heisenberg picture," where operators depend on time and states don't, then ϕ and π *do* depend on time. The natural question to ask, then, is *which* Hamiltonian we are using to introduce this time dependence.

We will, in fact, have occasion to use both, so let's introduce yet more notation to distinguish between the two possibilities: for a spacetime point $x = (t, \mathbf{x})$, let's write

$$\phi_0(x) = e^{iH_0 t} \phi(\mathbf{x}) e^{-iH_0 t}, \quad \phi(x) = e^{iH t} \phi(\mathbf{x}) e^{-iH t}.$$

In particular, these two versions of the field operator agree at $t = 0$ (that is, $\phi_0(0, \mathbf{x}) = \phi(0, \mathbf{x}) = \phi(\mathbf{x})$) but not at other times. It will be convenient to also have a name for the operator that translates between ϕ_0 and ϕ , so we'll also write

$$V(t, t') = e^{iH_0 t} e^{-iH(t-t')} e^{-iH_0 t'},$$

so that

$$\phi_0(x) = V(t, 0) \phi(x) V(0, t).$$

Note also that, with this definition, $V(t, t') V(t', t'') = V(t, t'')$.

Each $\phi(x)$ appearing in the time-ordered product we're trying to compute uses the *interacting* time evolution. Since the free Hamiltonian is much easier to understand than the interacting Hamiltonian, we'd like to rewrite everything in terms of states and operators from the free theory. As a first step, if the points x_1, \dots, x_n are in time order, I encourage you to show that we have

$$\begin{aligned} \langle \Omega | \phi(x_1) \cdots \phi(x_n) | \Omega \rangle \\ = \langle \Omega | V(0, t_1) \phi_0(x_1) V(t_1, t_2) \phi_0(x_2) V(t_2, t_3) \cdots V(t_{n-1}, t_n) \phi_0(x_n) V(t_n, 0) | \Omega \rangle. \end{aligned}$$

4.2 Isolating the Vacuum

Having successfully turned our ϕ 's into ϕ_0 's, let's now try to relate our two vacuum states $|\Omega\rangle$ and $|0\rangle$.

Let's write E_Ω for the eigenvalue of H associated to $|\Omega\rangle$. (Often one subtracts a constant from the Hamiltonian to make $E_\Omega = 0$. Since this doesn't affect any quantum-mechanical computations, you should feel free to either do this or not, and we're therefore going to be a bit cavalier about this detail.) By assumption, E_Ω is strictly smaller than all other eigenvalues of H .

Let $|\psi\rangle$ be any state satisfying $\langle\Omega|\psi\rangle \neq 0$, and imagine for the moment that the spectrum of H is discrete. (It's not, of course; we'll discuss that issue in just a moment.) We could then write

$$H|\psi\rangle = E_\Omega|\Omega\rangle\langle\Omega|\psi\rangle + \sum_{i=1}^{\infty} E_i|\psi_i\rangle\langle\psi_i|\psi\rangle,$$

where the $|\psi_i\rangle$'s are all the other eigenstates of H and the E_i 's are the associated eigenvalues.

For any T , we then have

$$e^{-iT(H-E_\Omega)}|\psi\rangle = |\Omega\rangle\langle\Omega|\psi\rangle + \sum_{i=1}^{\infty} e^{-iT(E_i-E_\Omega)}|\psi_i\rangle\langle\psi_i|\psi\rangle,$$

and we can isolate the first term with the following trick: let T go to infinity, but with a tiny negative imaginary part. Since $E_i - E_\Omega$ is strictly positive, we see that

$$\lim_{T \rightarrow \infty} e^{-iT(1-i\epsilon)(E_i-E_\Omega)} = 0.$$

Therefore,

$$|\Omega\rangle = \lim_{T \rightarrow \infty(1-i\epsilon)} \frac{e^{-iT(H-E_\Omega)}|\psi\rangle}{\langle\Omega|\psi\rangle}.$$

(From now on, rather than multiply T by $1 - i\epsilon$ every time it appears, we'll denote it in the limit like this.)

The spectrum of H is not discrete, but taking that fact into account only requires making a small modification to the argument: the sum should be replaced with a combination of sums and integrals to cover both the discrete and continuous parts of the spectrum. (The "correct" way to express this is in terms of a projection-valued measure, but I'm deliberately avoiding using that machinery here.) I encourage you to convince yourself that this doesn't affect the final conclusion.

If we take $|\psi\rangle$ to be the free vacuum $|0\rangle$ and write E_0 for the associated eigenvalue of H_0 , you can check that this equation becomes

$$|\Omega\rangle = \lim_{T \rightarrow \infty(1-i\epsilon)} \frac{e^{i(E_\Omega-E_0)T}V(0, -T)|0\rangle}{\langle\Omega|0\rangle}.$$

(At least, this will work if we have nonzero overlap between our two vacuum states! More on this assumption in the final section.) An exactly analogous argument shows that

$$\langle\Omega| = \lim_{T \rightarrow \infty(1-i\epsilon)} \frac{e^{i(E_\Omega-E_0)T}\langle 0|V(T, 0)}{\langle 0|\Omega\rangle}.$$

Plugging this into our earlier formula (remembering that we're still assuming that the x_i 's are time-ordered) gives

$$\begin{aligned} & \langle\Omega|\phi(x_1) \cdots \phi(x_n)|\Omega\rangle \\ &= \lim_{T \rightarrow \infty(1-i\epsilon)} \frac{e^{2i(E_\Omega-E_0)T}}{|\langle\Omega|0\rangle|^2} \langle 0|V(T, t_1)\phi_0(x_1)V(t_1, t_2) \cdots V(t_{n-1}, t_n)\phi_0(t_n)V(t_n, -T)|0\rangle. \end{aligned}$$

We can get rid of that annoying factor in front if we note that this formula also holds when there are zero field operators, meaning that

$$1 = \langle\Omega|\Omega\rangle = \lim_{T \rightarrow \infty(1-i\epsilon)} \frac{e^{2i(E_\Omega-E_0)T}}{|\langle\Omega|0\rangle|^2} \langle 0|V(T, -T)|0\rangle.$$

Using this, we get

$$\begin{aligned} & \langle \Omega | \phi(x_1) \cdots \phi(x_n) | \Omega \rangle \\ &= \lim_{T \rightarrow \infty (1-i\epsilon)} \frac{\langle 0 | V(T, t_1) \phi_0(x_1) V(t_1, t_2) \cdots V(t_{n-1}, t_n) \phi_0(t_n) V(t_n, -T) | 0 \rangle}{\langle 0 | V(T, -T) | 0 \rangle}. \end{aligned}$$

4.3 The Dyson Series

This is big progress! We'll be in business if we can express the operators $V(t, t')$ in terms of quantities that depend only on the free theory.

To accomplish this, we'll first note that

$$\frac{d}{dt} V(t, t') = \frac{d}{dt} (e^{iH_0 t} e^{-iH(t-t')} e^{-iH_0 t'}) = e^{iH_0 t} (-i(H - H_0)) e^{-iH_0 t} V(t, t').$$

Recalling that we defined the “interaction part” of the Hamiltonian as $H_I = H - H_0$, let's write

$$H_0^I(t) = e^{iH_0 t} H_I e^{-iH_0 t}$$

(by analogy with ϕ_0) so that this differential equation becomes simply

$$\frac{d}{dt} V(t, t') = -i H_0^I(t) V(t, t').$$

This, combined with the initial condition $V(t', t') = 1$, implies that this operator satisfies the integral equation

$$V(t, t') = 1 - i \int_{t'}^t H_0^I(\tau) V(\tau, t') d\tau.$$

Assuming everything converges nicely (which we'll hold off on worrying about for now) applying this formula recursively results in

$$V(t, t') = \sum_{n=0}^{\infty} (-i)^n \int_{t'}^t \int_{t'}^{\tau_n} \cdots \int_{t'}^{\tau_2} H_0^I(\tau_n) H_0^I(\tau_{n-1}) \cdots H_0^I(\tau_1) d\tau_1 \cdots d\tau_n.$$

Because of the limits of each of the integrals, the τ_i 's all appear in decreasing time order as long as $t \geq t'$. We can take advantage of this fact to write the series in a slightly more convenient way: if $t \geq t'$, then I encourage you to convince yourself that this series can be rewritten in the form

$$V(t, t') = \sum_{n=0}^{\infty} \frac{(-i)^n}{n!} \int_{t'}^t \cdots \int_{t'}^t T[H_0^I(\tau_n) \cdots H_0^I(\tau_1)] d\tau_1 \cdots d\tau_n.$$

This is called the **Dyson series** for $V(t, t')$. Because of how strongly this resembles the power series for an exponential function, it's very common use the suggestive notation

$$V(t, t') = T \exp \left[-i \int_{t'}^t H_0^I(\tau) d\tau \right],$$

but it's important to remember that this is just a convenient way to denote the previous expression, not a new mathematical fact.

This is a very encouraging step: we've managed to express $V(t, t')$ in terms of sums of integrals of $H_0^I(\tau)$, which in our case is equal to $\frac{\lambda}{4!} \int d^3\mathbf{x} \phi_0(\tau, \mathbf{x})^4$ — in particular, it involves

only free field operators, as desired. The rest is essentially bookkeeping. Recall that we are trying to compute

$$\lim_{T \rightarrow \infty(1-i\epsilon)} \frac{\langle 0|V(T, t_1)\phi_0(x_1)V(t_1, t_2)\cdots V(t_{n-1}, t_n)\phi_0(t_n)V(t_n, -T)|0\rangle}{\langle 0|V(T, -T)|0\rangle}.$$

For sufficiently large T (provided we use the real part of T to place it in time order) the factors appearing here are time-ordered, and in particular the first argument to each V is always greater than or equal to the second, so we are free to plug in the Dyson series. The numerator in this expression will look like

$$\langle 0|\text{Texp}\left[-i\int_{t_1}^T H_0^I(\tau)d\tau\right]\phi_0(x_1)\text{Texp}\left[-i\int_{t_2}^{t_1} H_0^I(\tau)d\tau\right]\phi_0(x_2)\cdots|0\rangle.$$

I encourage you to check that, when the time coordinates all appear in decreasing order as they do here, the Texp's take addition to multiplication in the same way exp's would, and so we can combine everything in this expression into one big time-ordering symbol, giving us

$$\langle 0|T\left[\phi_0(x_1)\cdots\phi_0(x_n)\exp\left(-i\int_{-T}^T H_0^I(\tau)d\tau\right)\right]|0\rangle.$$

(This is a less trivial statement than the Texp notation might suggest! I strongly urge you to work it out if it's at all unclear.) To be clear, the presence of the exp inside the time-ordering symbol means to expand exp in a power series and time-order each term along with the $\phi(x_i)$'s, that is, it means

$$\sum_{n=0}^{\infty}\langle 0|T\left[\phi_0(x_1)\cdots\phi_0(x_n)\frac{(-i)^n}{n!}\left(\int_{-T}^T H_0^I(\tau)d\tau\right)^n\right]|0\rangle.$$

Plugging this into our equation for the time-ordered n -point function from earlier, we finally get our result:

$$\langle \Omega|T[\phi(x_1)\cdots\phi(x_n)]|\Omega\rangle = \lim_{T \rightarrow \infty(1-i\epsilon)} \frac{\langle 0|T\left[\phi_0(x_1)\cdots\phi_0(x_n)\exp\left(-i\int_{-T}^T H_0^I(\tau)d\tau\right)\right]|0\rangle}{\langle 0|\text{Texp}\left[-i\int_{-T}^T H_0^I(\tau)d\tau\right]|0\rangle}.$$

5 Wick's Theorem and Feynman Diagrams

5.1 Wick's Theorem

We've now managed to reduce everything to the question of how to evaluate the expectation value of a time-ordered product of field operators in the *free* theory. Because all the computations we have left only involve free-theory operators, we'll drop all the subscript-0's, so that the expressions we're interested in evaluating has the form

$$\langle 0|T[\phi(z_1)\cdots\phi(z_m)]|0\rangle.$$

We'll start by computing this, and then plug the result into the Dyson series. The computations we're about to do are covered very well in the physics textbooks, and there aren't really any conceptual issues worth digging into until later in the story, so while I'll try to state everything

clearly, I'm also going to be quite brief. For more details, I recommend the discussions in the books mentioned in the introduction. In particular, this is done in Sections 4.2–4.3 of Peskin and Schroeder, and in Section 6.4 of Folland.

In the free theory, we can expand the field operator $\phi(x)$ in terms of creation and annihilation operators. Specifically, let's write

$$\phi(x) = \phi^+(x) + \phi^-(x),$$

where

$$\phi^+(x) = \int \frac{d^3\mathbf{p}}{(2\pi)^{3/2}\sqrt{2\omega_{\mathbf{p}}}} a(\mathbf{p})e^{-ip \cdot x}, \quad \phi^-(x) = \int \frac{d^3\mathbf{p}}{(2\pi)^{3/2}\sqrt{2\omega_{\mathbf{p}}}} a^\dagger(\mathbf{p})e^{ip \cdot x},$$

and where as always in expressions like this the time component of p is taken to be $\omega_{\mathbf{p}}$.

The reason to split up $\phi(x)$ in this way is that both $\phi^+(x)|0\rangle$ and $\langle 0|\phi^-(x)$ are zero, and so we might be able to get a handle on our computation if we can find a way to rewrite any product of $\phi(x)$'s in terms of an expression which pushes all of the $\phi^+(x)$'s to the right and all of the $\phi^-(x)$'s to the left. I encourage you use the commutation relation $[a(\mathbf{p}), a^\dagger(\mathbf{p}')] = \delta(\mathbf{p} - \mathbf{p}')$ to verify that

$$[\phi^+(x), \phi^-(y)] = \int \frac{d^3\mathbf{p}}{(2\pi)^3(2\omega_{\mathbf{p}})} e^{-ip \cdot (x-y)}.$$

We'll call this last function $\Delta^m(x - y)$. (It depends on m through its dependence on $\omega_{\mathbf{p}} = \sqrt{|\mathbf{p}|^2 + m^2}$.)

Let's start with the case where there are just two field operators, that is, $\langle 0|T[\phi(x)\phi(y)]|0\rangle$. If we write x^0 and y^0 for the time components of x and y then, if $x^0 > y^0$, we have

$$\begin{aligned} \langle 0|T[\phi(x)\phi(y)]|0\rangle &= \langle 0|\phi(x)\phi(y)|0\rangle \\ &= \langle 0|\phi^+(x)\phi^+(y) + \phi^-(x)\phi^+(y) + \phi^+(x)\phi^-(y) + \phi^-(x)\phi^-(y)|0\rangle \\ &= \langle 0|\phi^+(x)\phi^+(y) + \phi^-(x)\phi^+(y) + \phi^-(y)\phi^-(x) + \phi^-(x)\phi^-(y)|0\rangle + [\phi^+(x), \phi^-(y)] \\ &= \Delta^m(x - y), \end{aligned}$$

because, on the second-to-last line, every term inside the expectation value either has a ϕ^+ next to $|0\rangle$ or a ϕ^- next to $\langle 0|$. On the other hand, if $x^0 \leq y^0$, then we get $\Delta^m(y - x)$ for the same reason. In other words,

$$\langle 0|T[\phi(x)\phi(y)]|0\rangle = \begin{cases} \Delta^m(x - y) & x^0 > y^0 \\ \Delta^m(y - x) & x^0 \leq y^0 \end{cases}.$$

This function is important enough to have a name: it's written $\Delta_F^m(x - y)$, and it's called the **Feynman propagator**.

It is hopefully easy to believe — and I encourage you to prove it to yourself if you're so inclined — that when we have a larger number of field operators we can apply this procedure repeatedly to turn the entire expression into a product of propagators. The final result is called **Wick's Theorem**, and it states that $\langle 0|T[\phi(z_1) \cdots \phi(z_k)]|0\rangle$ is zero if k is odd, and is equal to

$$\sum_{(u_1, v_1), \dots, (u_{k/2}, v_{k/2})} \prod_{i=1}^{k/2} \Delta_F^m(u_i - v_i)$$

if m is even, where the sum ranges over all $(k - 1)!!$ ways of pairing off the z_j 's into $k/2$ pairs. (Notice that Δ_F^m is an even function, so it doesn't matter how we order the two points within each pair (u_i, v_i) .) For example, when $k = 4$, the theorem says that

$$\begin{aligned} \langle 0|T[\phi(z_1)\phi(z_2)\phi(z_3)\phi(z_4)]|0\rangle \\ = \Delta_F^m(z_1 - z_2)\Delta_F^m(z_3 - z_4) + \Delta_F^m(z_1 - z_3)\Delta_F^m(z_2 - z_4) + \Delta_F^m(z_1 - z_4)\Delta_F^m(z_2 - z_3). \end{aligned}$$

5.2 Feynman Diagrams in Position Space

Now let's see how this affects the expression we're trying to evaluate which, I will remind you, looks like

$$\lim_{T \rightarrow \infty(1-i\epsilon)} \frac{\langle 0|T\left[\phi(x_1)\cdots\phi(x_n)\exp\left(-i\int_{-T}^T H_I(\tau)d\tau\right)\right]|0\rangle}{\langle 0|T\exp\left[-i\int_{-T}^T H_I(\tau)d\tau\right]|0\rangle}.$$

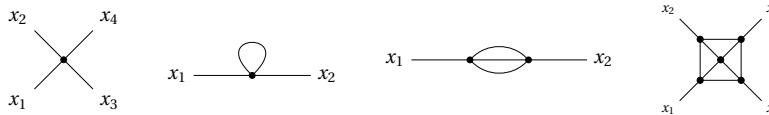
We'll focus on the numerator for now. If we plug in the definition of H_I and expand the exp as a power series, it looks like

$$\begin{aligned} \lim_{T \rightarrow \infty(1-i\epsilon)} \sum_{k=0}^{\infty} \frac{1}{k!} \left(\frac{-i\lambda}{4!}\right)^k \\ \times \int_{-T}^T d\tau_1 \cdots \int_{-T}^T d\tau_k \int d^3\mathbf{y}_1 \cdots \int d^3\mathbf{y}_k \langle 0|T[\phi(x_1)\cdots\phi(x_n)\phi(y_1)^4\cdots\phi(y_k)^4]|0\rangle, \end{aligned}$$

where we write $y_i = (\tau_i, \mathbf{y}_i)$. The integrand then has exactly the form we just computed, and so we can use Wick's Theorem to replace it with an appropriate sum of products of propagators. Each term in the sum will correspond to one way of splitting the list of variables $(x_1, \dots, x_n, y_1, \dots, y_k)$ into pairs, where each x_i appears once and each y_i appears four times.

There is a very convenient way to represent the resulting product of propagators as a graph: we can denote each variable as a vertex, and each propagator $\Delta_F^m(z_i - z_j)$ as an edge connecting the z_i vertex with the z_j vertex. (Because Δ_F^m is an even function, the order of z_i and z_j doesn't matter, so this is an *undirected* edge.) Because each x_i appears once and each y_i appears four times, the x_i vertices will each be attached to just one edge and the y_i vertices will each be attached to four edges. We'll call the former **external vertices** and the latter **internal vertices**. The external vertices are all labeled with the corresponding position variable x_i , but the internal vertices are by convention unlabeled.

The resulting graphs are called **position-space Feynman diagrams**, and as you're probably aware, Feynman diagrams are absolutely everywhere in perturbative quantum field theory. Here are a few examples, with position labels on the external vertex and with the internal vertices marked with solid dots:



Feynman diagrams are much more convenient to work with than the expressions we've been manipulating up to this point, so it would be wonderful if we could completely forget about integrals of time-ordered products of field operators and work directly with the diagrams. This

is basically where we will end up, but there is one more detail that needs to be handled before that transition can be complete: we need to determine how many terms of the sum will produce the same diagram. If we can manage to compute this number just by looking at the diagram, then we can just multiply the value associated to that diagram by that number and sum over all diagrams to get our result.

There are two straightforward ways that one diagram will correspond to many different terms. One arises from the fact that the internal vertices of a diagram, which correspond to the y_i 's, are unlabeled. Since we can permute the y_i 's any way we want without changing the diagram, each diagram will, in general, correspond to $k!$ different terms. Luckily, there is a factor of $1/k!$ ready and waiting in front of our integral to cancel it.

Similarly, each y_i appears four times, so if a given internal vertex is paired with four distinct other vertices, we can also permute the four copies of y_i without changing the diagram, which multiplies the number of terms corresponding to the diagram by an additional factor of $4!$. This contribution is also nicely cancelled by the factor of $1/4!$ attached to each of the k copies of λ in front of the integral. (In fact, it was for precisely this reason that we included that factor in the Hamiltonian in the first place.)

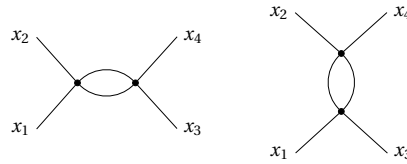
There are, though, a couple cases where the situation is not quite so simple. Consider the third diagram above, with two internal vertices we can call y_1 and y_2 connected by three edges, with y_1 also connected to the external vertex x_1 and y_2 also connected to the external vertex x_2 . I encourage you to check that there are only $4 \cdot 4!$ ways to assign the four copies of y_1 and y_2 to their respective pairs, not the $(4!)^2$ that the analysis from the previous paragraph would suggest. Therefore, after we divide by the two copies of $4!$, this diagram ends up with a factor of $1/3!$.

This $3!$ is called a **symmetry factor**, and keeping track of these is one of the more annoying aspects of the Feynman diagram story. Because our focus in this series is on conceptual understanding rather than the details of the computations, I'm not going to give a complete theory of symmetry factors here. If you are ever unsure, you can always go back to the product of propagators that arises from Wick's theorem.

Here the products of propagators we get from the first three examples above, with the symmetry factors included (to cut down on notation, we've written the integrals over the internal position variables as $\int d^4y$, without splitting up the time and space parts and without the $\lim_{T \rightarrow \infty} (1 - i\epsilon)$ in front):

$$\begin{array}{l}
 \begin{array}{c} x_2 \\ \diagdown \\ \bullet \\ \diagup \\ x_1 \end{array} \begin{array}{c} x_4 \\ \diagdown \\ \bullet \\ \diagup \\ x_3 \end{array} \quad -i\lambda \int d^4y \Delta_F^m(x_1 - y) \Delta_F^m(x_2 - y) \Delta_F^m(x_3 - y) \Delta_F^m(x_4 - y) \\
 \\
 x_1 \text{ --- } \text{loop} \text{ --- } x_2 \quad \frac{-i\lambda}{2} \int d^4y \Delta_F^m(x_1 - y) \Delta_F^m(x_2 - y) \Delta_F^m(0) \\
 \\
 x_1 \text{ --- } \text{double loop} \text{ --- } x_2 \quad \frac{(-i\lambda)^2}{6} \int d^4y_1 \int d^4y_2 \Delta_F^m(x_1 - y_1) \Delta_F^m(x_2 - y_2) \Delta_F^m(y_1 - y_2)^3
 \end{array}$$

Two Feynman diagrams are considered “the same” if there is a bijection between the vertices and edges of one diagram and the vertices and edges of the other diagram that preserves the labels of all the external vertices. For example, in the first example above with four external vertices all connected to a single internal vertex, any permutation of the external labels will result in the same diagram, whereas these two diagrams are different:



With this last detail in place, the story so far can be summarized in a quite pleasing way. To compute the m 'th term in the power series representation for our numerator

$$\langle 0|T \left[\phi(x_1) \cdots \phi(x_n) \exp \left(-i \int_{-T}^T H_I(\tau) d\tau \right) \right] |0\rangle,$$

first draw all possible Feynman diagrams with n external (that is, degree-one) vertices and m internal (degree-four) vertices, where the external vertices are labeled with the x_i 's. We can assign a value to each diagram as follows:

- Introduce new variables $y_i = (\tau_i, \mathbf{y}_i)$ for each internal vertex.
- For each internal vertex $y = (\tau, \mathbf{y})$, include a factor of $-i\lambda$ and an integral $\int_{-T}^T d\tau \int d^3\mathbf{y}$.
- For each edge connecting a vertex u to a vertex v , include a factor of $\Delta_F^m(u - v)$.
- Take the limit $T \rightarrow \infty(1 - i\epsilon)$ and divide by the symmetry factor.

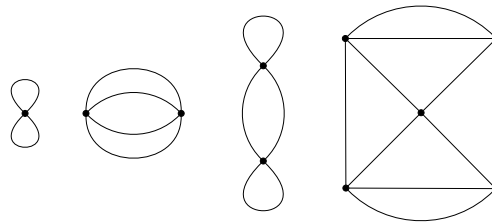
These are called the **position-space Feynman rules** for the interacting quantum field theory we're evaluating. There are, of course, infinitely many Feynman diagrams, but only finitely many arising from any given term in the Dyson series. We therefore have an algorithm for computing our power series to any degree we might want, or as physicists say, to any desired "degree in perturbation theory."

5.3 Vacuum Bubbles

This handles the numerator of our fraction, but we also have the denominator, which I will remind you looked like

$$\langle 0|T \exp \left[-i \int_{-T}^T H_I(\tau) d\tau \right] |0\rangle.$$

The Feynman rules we just derived in fact also apply equally well to this expression: this is simply the special case where there are no external vertices. Diagrams with no external vertices are called **vacuum bubbles**. Here are a couple examples:



Let's look at the first of these diagrams and try to assign a value to it. According to the Feynman rules, we should assign it the value

$$\lim_{T \rightarrow \infty(1-i\epsilon)} \frac{-i\lambda}{8} \int_{-T}^T d\tau \int d^3\mathbf{y} \Delta_F^m(0)^2.$$

This integral is divergent several times over: $\Delta_F^m(0)$ is infinite, and even if it weren't, we would be integrating a constant over all of \mathbb{R}^4 .

In fact, essentially the same thing will happen with *any* vacuum bubble diagram. (This will probably be easier to prove to yourself after we've switched to the momentum-space version of the Feynman rules in the next subsection.) The problem is not limited to the denominator: it's a simple consequence of the Feynman rules that whenever a diagram D is disconnected, say as the disjoint union of diagrams D_1 and D_2 , then the value of D is the product of the values of D_1 and D_2 , which means that any diagram that *contains* a vacuum bubble as a connected component will also diverge.

If you're willing to think like a physicist and play a little fast and loose with infinities, there's a surprisingly simple fix. Any Feynman diagram with n external vertices can be written as a disjoint union of a (possibly empty) vacuum bubble and a diagram with no vacuum bubbles. So, by our disjoint-unions-go-to-products rule, since our numerator is the sum of the values of all such diagrams, we can write the numerator as a product of two sums: it's the sum over all vacuum bubbles times the sum over all diagrams *without* any vacuum bubbles.

The first of these factors then cancels with the denominator. Therefore, at least to the extent you buy this argument, we've arrived at a very elegant way to describe the entire fraction

$$\lim_{T \rightarrow \infty(1-i\epsilon)} \frac{\langle 0|T \left[\phi(x_1) \cdots \phi(x_n) \exp \left(-i \int_{-T}^T H_I(\tau) d\tau \right) \right] |0\rangle}{\langle 0|T \exp \left[-i \int_{-T}^T H_I(\tau) d\tau \right] |0\rangle}$$

as the power series whose k 'th term is simply the sum of the values of all Feynman diagrams with k internal vertices *and no vacuum bubbles*.

You could be excused for finding this argument unsatisfying; it does, after all, rely on cancelling infinity with infinity and assuming you get 1. We'll be in a much better place to describe what's going on here after we've discussed renormalization in the next article, but since it's easy to get stuck on points like this I want to say a few words about the shape that story will take.

This is, in fact, far from the last time we'll see a divergent integral emerging from a Feynman diagram. In those cases as well as this one, we will attempt to resolve the issue by writing all of our Feynman integrals as limits of some finite quantities (called *regularized* integrals) as some parameter (called a *cutoff*) goes to infinity. We will then be in a position to perform whatever cancellations need to happen, like the step of dividing out the vacuum bubbles we just discussed, on the regularized quantities, *before* we send the cutoff to infinity. If we are sufficiently careful about how we juggle the relevant limits, it will in fact be possible to extract actual, finite numbers from this whole procedure that can be compared with experiment. We'll come back to this point in the final section.

5.4 Feynman Diagrams in Momentum Space

With the vacuum bubbles handled, we've accomplished the goal of expressing $G^{(n)}$ in terms of Feynman diagrams. But the quantity that appears in the LSZ formula is the Fourier transform,

$$\tilde{G}^{(n+m)}(p_1, \dots, p_n, -p_{n+1}, \dots, -p_{n+m}).$$

It turns out that there is a nice way to write this quantity where we integrate out the position variables corresponding to a diagram's internal vertices resulting in an expression where the only remaining variables are momenta. The resulting expression will be more straightforward to compute, and it will also tell us a bit more about how to interpret the value associated to a given diagram.

To accomplish this, it will help to take a more careful look at our expression for the Feynman propagator Δ_F^m . Recall that $\Delta_F^m(x)$ is equal to either $\Delta^m(x)$ or $\Delta^m(-x)$, depending on whether the time component of x is larger or smaller than 0. Using the fact that

$$\Delta^m(x) = \int \frac{d^3\mathbf{p}}{(2\pi)^3(2\omega_{\mathbf{p}})} e^{-ip \cdot x},$$

I encourage you to verify that this means that

$$\Delta_F^m(x) = \int \frac{d^3\mathbf{p}}{(2\pi)^3(2\omega_{\mathbf{p}})} e^{-i\omega_{\mathbf{p}}|t|} e^{i\mathbf{p} \cdot \mathbf{x}},$$

where $x = (t, \mathbf{x})$.

When Δ_F^m is written in this way, the only place that the spatial components of x appear is in the factor $e^{-i\mathbf{p} \cdot \mathbf{x}}$, and the result of integrating this over \mathbf{x} is simply $(2\pi)^3 \delta(\mathbf{p})$. In other words, this expression will already allow us to eliminate the spatial components \mathbf{y}_i of the position variables. In order to do the same for the time components τ_i , we want a way to introduce a similar integral over the time component of p .

In fact, the resulting expression is quite nice: we have

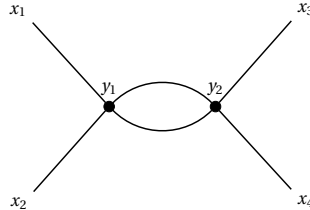
$$\Delta_F^m(x) = \lim_{\epsilon \rightarrow 0^+} \int \frac{d^4p}{(2\pi)^4} \frac{i}{p^2 - m^2 + i\epsilon} e^{-ip \cdot x}.$$

The computation is a fun exercise in complex analysis which I will mostly leave to you to work out, with just a few points to get you started. For very small positive values of ϵ , the integrand has two poles: one near $\omega_{\mathbf{p}}$ but just below the real axis, and one near $-\omega_{\mathbf{p}}$ but just above the real axis. To evaluate the integral over the time component of p , you can start with an integral from $-E$ to E along the real axis, complete it with a semicircle in either the upper or lower half plane depending on the sign of t , and then take E to ∞ .

Rather than sticking the $i\epsilon$ in the denominator like this, we could instead leave the denominator in the form $\frac{i}{p^2 - m^2}$ and slightly tilt the line we're integrating over counter-clockwise from the real axis. In other words, we could also say that

$$\Delta_F^m(x) = \lim_{\epsilon \rightarrow 0^+} \lim_{E \rightarrow \infty} \int_{-E}^E \frac{dp^0}{2\pi} \int \frac{d^3\mathbf{p}}{(2\pi)^3} \frac{i}{p^2 - m^2} e^{-ip \cdot x}.$$

To see how this affects our Feynman diagram integrals, it will be easier to see what's going on if we look at an example. Consider the following diagram, where I've also given names to the internal vertices this time for later convenience:



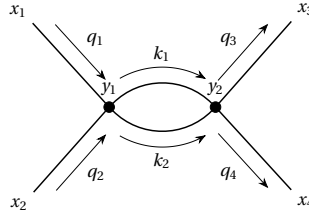
The diagram makes the following contribution to $\tilde{G}^{(4)}(p_1, p_2, -p_3, -p_4)$:

$$\begin{aligned} & \frac{(-i\lambda)^2}{2} \int d^4x_1 \int d^4x_2 \int d^4x_3 \int d^4x_4 \int d^4y_1 \int d^4y_2 e^{-i(p_1 \cdot x_1 + p_2 \cdot x_2 - p_3 \cdot x_3 - p_4 \cdot x_4)} \\ & \quad \times \Delta_F^m(y_1 - x_1) \Delta_F^m(y_1 - x_2) \Delta_F^m(y_2 - y_1)^2 \Delta_F^m(x_3 - y_2) \Delta_F^m(x_4 - y_2) \end{aligned}$$

(To save on notation, we're writing the integrals over the internal position variables in the form $\int d^4y$ rather than $\lim_{T \rightarrow \infty(1-i\epsilon)} \int_{-T}^T d\tau \int d^3\mathbf{y}$.)

If we replace each copy of Δ_F^m with the expression we just derived, this has the effect of introducing a new momentum variable for every edge in the diagram. Notice that, even though $\Delta_F^m(u - v)$ is symmetric in u and v , the integrand $\frac{i}{p^2 - m^2 + i\epsilon} e^{-ip \cdot (u-v)}$ isn't, so when we attach a momentum label to an edge we need to also arbitrarily assign the edge a direction to go along with it: let's adopt the convention that, when an edge is labeled p and points from u to v , that means we have a factor of $e^{-ip \cdot (v-u)}$ in the integral. When we do this, we'll say that the momentum "flowing out of u " or "into v " along this edge is p ; equivalently, we could say that the momentum "flowing into u " or "out of v " is $-p$.

After adorning our diagram with momentum labels, it might look like this:



And our integral — don't worry, it's about to get a lot simpler! — now looks like this:

$$\begin{aligned} & \frac{(-i\lambda)^2}{2} \int d^4x_1 \int d^4x_2 \int d^4x_3 \int d^4x_4 \int d^4y_1 \int d^4y_2 \\ & \quad \times \int \frac{d^4q_1}{(2\pi)^4} \int \frac{d^4q_2}{(2\pi)^4} \int \frac{d^4q_3}{(2\pi)^4} \int \frac{d^4q_4}{(2\pi)^4} \int \frac{d^4k_1}{(2\pi)^4} \int \frac{d^4k_2}{(2\pi)^4} \\ & \quad \times e^{-i(p_1 \cdot x_1 + p_2 \cdot x_2 - p_3 \cdot x_3 - p_4 \cdot x_4)} e^{-i(q_1 \cdot (y_1 - x_1) + q_2 \cdot (y_1 - x_2) + k_1 \cdot (y_2 - y_1) + k_2 \cdot (y_2 - y_1) + q_3 \cdot (x_3 - y_2) + q_4 \cdot (x_4 - y_2))} \\ & \quad \times \frac{i}{q_1^2 - m^2} \frac{i}{q_2^2 - m^2} \frac{i}{q_3^2 - m^2} \frac{i}{q_4^2 - m^2} \frac{i}{k_1^2 - m^2} \frac{i}{k_2^2 - m^2} \end{aligned}$$

(To save on notation, we are again not writing the $\lim_{T \rightarrow \infty(1-i\epsilon)}$ for the y_i integrals, and we're also not including the $+i\epsilon$ in the denominators of the propagators. We'll come back to the effect of these two $i\epsilon$'s in just a moment.)

The thing to notice here is that the position variables, whether internal or external, now appear only in the exponentials on the third line of this giant expression. Specifically, the part of the integrand involving an internal variable y_i will look like $e^{-i(k_1 + k_2 + k_3 + k_4) \cdot y_i}$, where the k_j 's are the momenta flowing *into* the vertex. (Remember that this means that if our chosen orientation has one of the momenta flowing *out of* the vertex we flip the sign!) Evaluating the integral over y_i will then just produce $(2\pi)^4 \delta(k_1 + k_2 + k_3 + k_4)$.

For the external vertices x_i — provided that, as in our example, we oriented the edges so that momentum flows into the diagram from incoming particles and out of the diagram from outgoing particles — we simply end up with a $(2\pi)^4 \delta(p_i - q_i)$. This can then be combined with

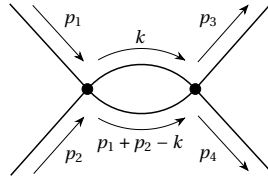
the integral over q_i , which takes the form $\int \frac{d^4 q_i}{(2\pi)^4}$, the net effect of which is to simply identify each p_i with the corresponding q_i .

When we do these $e^{-ik \cdot y}$ integrals, there is a happy collision between the two $(1 \pm i\epsilon)$'s that appear in our spacetime and momentum integrals. It's easier to see if we use the second of the two expressions for Δ_F^m we derived above, the one with the $\lim_{E \rightarrow \infty(1+i\epsilon)}$. Then, in the T and E limits, the time component of y_i is multiplied by $1 - i\epsilon$, and the time components of the momenta are multiplied by $1 + i\epsilon$. When we multiply them, if we take our limits carefully enough, the small clockwise rotation on the position variables will cancel the small counter-clockwise rotation on the momentum variables. This is fortunate: if we had only one of these but not the other, then the integral would blow up for either large positive or large negative values of τ_i .

By evaluating all of these integrals, we can successfully eliminate all of the position variables. In fact, we can even go further: by repeatedly applying the fact

$$\int \frac{d^4 b}{(2\pi)^4} [(2\pi)^4 \delta(a - b)] [(2\pi)^4 \delta(b - c)] = (2\pi)^4 \delta(a - c),$$

we can use the resulting delta functions to eliminate many of the momentum variables as well. It's customary to write the momentum labels on a diagram with as many of these delta functions integrated out as possible; in our running example, that might look like this:



The resulting integral, representing the contribution of this diagram to $\tilde{G}^{(4)}(p_1, p_2, -p_3, -p_4)$ is then

$$\lim_{\epsilon \rightarrow 0^+} \frac{(-i\lambda)^2}{2} (2\pi)^4 \delta(p_1 + p_2 - p_3 - p_4) \frac{i}{p_1^2 - m^2 + i\epsilon} \frac{i}{p_2^2 - m^2 + i\epsilon} \frac{i}{p_3^2 - m^2 + i\epsilon} \frac{i}{p_4^2 - m^2 + i\epsilon} \\ \times \int \frac{d^4 k}{(2\pi)^4} \frac{i}{k^2 - m^2 + i\epsilon} \frac{i}{(p_1 + p_2 - k)^2 - m^2 + i\epsilon}.$$

In general, we'll be left with a number of momentum integrals equal to the first Betti number of the Feynman diagram, what physicists refer to as the number of **loops** in the diagram. As long as the diagram is connected (which we'll assume for simplicity), we can collect all of the delta functions into a single "global momentum conservation" delta function involving the incoming and outgoing momenta.

We can summarize the result of this whole procedure in a series of **momentum-space Feynman rules**. To write the m 'th term in our perturbation series, draw all possible Feynman diagrams with no vacuum bubbles and with n external vertices and m internal vertices, where the edges connecting to external vertices are labeled with the p_i 's so that incoming momenta point into the diagram and outgoing momenta point out of it. Then each diagram is assigned a value as follows:

- Introduce new variables k_i for each internal edge, orienting them however you want.
- For each internal vertex, include a factor of $-i\lambda(2\pi)^4 \delta(k_1 + k_2 + k_3 + k_4)$, where the sum is over all momenta flowing *into* the vertex.

- For each edge k , introduce a factor of $\frac{i}{k^2 - m^2 + i\epsilon}$ and an integral $\int \frac{d^4 k}{(2\pi)^4}$.
- Take the limit $\lim_{\epsilon \rightarrow 0^+}$ and divide by the symmetry factor.

For example, here are the diagrams we considered earlier when discussing the position-space rules along with their contributions to $\tilde{G}^{(n)}$ according to the momentum-space rules (without the $+i\epsilon$'s in the propagators):

$$\begin{aligned}
 & -i\lambda(2\pi)^4 \delta(p_1 + p_2 - p_3 - p_4) \frac{i}{p_1^2 - m^2} \frac{i}{p_2^2 - m^2} \frac{i}{p_3^2 - m^2} \frac{i}{p_4^2 - m^2} \\
 & \frac{-i\lambda}{2} (2\pi)^4 \delta(p_1 - p_2) \frac{i}{p_1^2 - m^2} \frac{i}{p_2^2 - m^2} \int \frac{d^4 k}{(2\pi)^4} \frac{i}{k^2 - m^2} \\
 & \frac{(-i\lambda)^2}{6} (2\pi)^4 \delta(p_1 - p_2) \frac{i}{p_1^2 - m^2} \frac{i}{p_2^2 - m^2} \int \frac{d^4 k_1}{(2\pi)^4} \int \frac{d^4 k_2}{(2\pi)^4} \frac{i}{k_1^2 - m^2} \frac{i}{k_2^2 - m^2} \frac{i}{(p_1 - k_1 - k_2)^2 - m^2}
 \end{aligned}$$

(You might notice that these integrals don't look especially convergent. If so, you're right! We'll come back to this in the final section.)

5.5 Interpreting Feynman Diagrams

Our Feynman diagram story is almost crying out for a physical interpretation: we can imagine n particles coming in from the left side of the diagram, sometimes “colliding” with other particles at the internal vertices, and ending up as m particles going out the right side. The internal edges of the diagram can then be said to describe the paths of so-called **virtual particles**. Each collision event can be thought of as two particles (real or virtual) hitting each other and leaving as two particles, or as one particle turning into three, or four into zero, etc. The delta functions attached to all of the internal vertices mean that each of these “collision events” has to conserve momentum — and therefore that momentum has to be conserved over the entire scattering process — which also serves to make this story more compelling.

At least at an informal level, this story can be a useful one to have in mind. But I think it's probably best not to take it too literally. For one thing, the momenta we're attaching to the “virtual particles” don't actually have to be valid momenta for a particle of mass m , and in fact they don't even have to be timelike. While you can try (and some authors do try) to tell a story where a particle is allowed to travel faster than the speed of light as long as it's “virtual,” one certainly has to admit that this puts the literal version of the story under quite a bit of strain.

I think it makes the most sense to just not try to treat Feynman diagrams as literal depictions of actual physical events involving particles. Remember where they came from: they arose as a convenient way to keep track of the terms in a power series expansion, in other words, as a bookkeeping device for performing a specific computation in perturbation theory. Opinions definitely differ on this point, but I find it easier to keep track of what's going on in quantum field theory when I think of the fields as the things the theory is most directly talking about and

particles as asymptotic descriptions of certain states that a collection of quantum fields might be in.

If you think of particles this way, as “excitations” in a quantum field, you absolve yourself of the obligation to answer questions like “how many particles are present at this exact moment in time?” Instead, a scattering process can be described by saying that a couple of particle-like excitations in the fields bumped into each other, the resulting not-very-particle-like state churned around for a while, and eventually some other once-again-particle-like excitations shot off from this blob of activity in various directions. In this interpretation, the Feynman diagrams are merely a convenient method for computing the probability amplitude for this process to end up a certain way.

With that said, the interpretation of Feynman diagrams in terms of virtual particles is so compelling that we will probably end up falling back on it ourselves as we progress through this series of articles. This is, in my opinion, fine as long as we remember not to take it too literally. As long as we keep this in mind, we’ll find that Feynman diagrams still have a lot to teach us.

6 The Mass and the Field Normalization

There is now just one last loose end that needs to be addressed. We’re trying to compute scattering amplitudes using the LSZ formula, which I’ll reproduce here for our convenience:

$$\langle p_{n+1}, \dots, p_{n+m} | p_1, \dots, p_n \rangle^{\text{out}} = \frac{(-i)^{n+m}}{Z^{(n+m)/2}} \prod_{k=1}^{n+m} (p_k^2 - \mu^2) \tilde{G}^{(n+m)}(p_1, \dots, p_n, -p_{n+1}, \dots, -p_{n+m}).$$

Our discussion so far has given us a way to compute $\tilde{G}^{(n+m)}$, at least perturbatively, which gets us most of the way toward an understanding on the right-hand side of this equation. But we don’t yet have any way to compute Z and μ . (In particular, recall that the mass μ of the particle is not the same as the parameter m in the Lagrangian!)

In the free theory, the mass of the particle can be extracted from the Feynman propagator: as we showed in the last section, we have

$$\langle 0 | T[\phi(x)\phi(y)] | 0 \rangle = \Delta_F^m(x-y) = \lim_{\epsilon \rightarrow 0^+} \int \frac{d^4 p}{(2\pi)^4} \frac{i}{p^2 - m^2 + i\epsilon} e^{-ip \cdot x}.$$

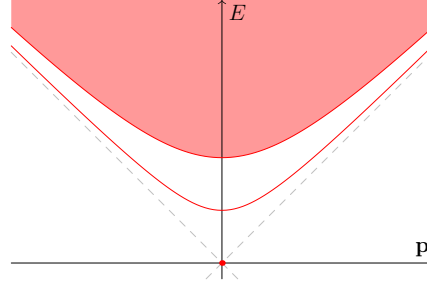
In other words, we can identify the mass of the particle in the free theory as the location of the pole in $\tilde{G}^{(2)}$.

Since we now have a handle on how to compute the $\tilde{G}^{(n)}$ ’s in the interacting theory, it makes sense to see if we can tell a similar story there. It might seem like we could learn about the two-point function by running the LSZ formula in reverse: since in and out states coincide for a single particle, it looks like LSZ would give us a relationship between $\langle q | p \rangle$ and $\tilde{G}^{(2)}(p, -q)$. Unfortunately, LSZ is only valid if the momenta involved are distinct, and in the case of one incoming and one outgoing particle, this means both sides of the equation are identically zero.

It would probably be possible to get what we need here by repeating the analysis that led to the LSZ formula while loosening the assumption that the momenta are distinct. But we’re going to establish the relationship between $\tilde{G}^{(2)}$ and particles in a different way that is both simpler and conceptually illuminating in its own right.

6.1 The Källén–Lehmann Spectral Representation

Let's start by recalling how the one-particle states fit into the Hilbert space. Recall that the components of the operator $P = (H, P_x, P_y, P_z) = (H, \mathbf{P})$ all commute with each other, which means they can be simultaneously diagonalized, and that each simultaneous eigenstate is also an eigenstate of $P^2 = H^2 - |\mathbf{P}|^2$. In the last article, we drew a picture of the spectrum of P in the free theory:



The red dot at the bottom is the vacuum state, the hyperbola in the middle is the one-particle states, and the region at the top accounts for all the multi-particle states. We've been assuming that this same picture continues to hold in our interacting theory. (At least, we're assuming this is true after subtracting a constant from Hamiltonian to make the energy of the vacuum equal to 0, which doesn't affect anything physically relevant; for the rest of this discussion, suppose we've done this.) In fact, if that hyperbola in the middle is $p^2 = \mu^2$, we *defined* the one-particle states as the eigenstates of P^2 with eigenvalue μ^2 .

We normalized the one-particle states by setting

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

Let's give names to the eigenstates belonging to the continuous part of the spectrum of P^2 and normalize them in a similar way: we'll write $|M_{\mathbf{p}}^2\rangle$ for the eigenstate whose eigenvalue under P^2 is M^2 and whose eigenvalues under $\mathbf{P} = (P_x, P_y, P_z)$ are the components of \mathbf{p} . We'll then set

$$\langle M_{\mathbf{p}'}^2 | M_{\mathbf{p}}^2 \rangle = (2\pi)^3 (2\omega_{\mathbf{p}}(M^2)) \delta(\mathbf{p}' - \mathbf{p}) \delta(M'^2 - M^2),$$

where $\omega_{\mathbf{p}}(M^2) = \sqrt{|\mathbf{p}|^2 + M^2}$. We'll also, in order to unify the notation, write $|\mu_{\mathbf{p}}^2\rangle$ for $|p\rangle$ for the duration of this argument.

With these conventions in place, we can expand an arbitrary state $|\psi\rangle$ in terms of eigenstates by writing

$$|\psi\rangle = |\Omega\rangle \langle \Omega | \psi \rangle + \int \frac{d^3\mathbf{p}}{(2\pi)^3 (2\omega_{\mathbf{p}})} |\mu_{\mathbf{p}}^2\rangle \langle \mu_{\mathbf{p}}^2 | \psi \rangle + \int_{M_0^2}^{\infty} d(M^2) \int \frac{d^3\mathbf{p}}{(2\pi)^3 (2\omega_{\mathbf{p}}(M))} |M_{\mathbf{p}}^2\rangle \langle M_{\mathbf{p}}^2 | \psi \rangle,$$

where M_0^2 is the value where the continuous part of the spectrum of P^2 starts. (You can verify that this is the right formula by plugging in our various eigenvectors for $|\psi\rangle$ and invoking the normalization conventions we just described! Also, note that we're following the physicists' convention of always referring to the eigenvalues as " M^2 ", hence the slightly odd-looking $d(M^2)$ in the integral.)

Let's stick this relation in between the two field operators in the expression $\langle \Omega | \phi(x) \phi(y) | \Omega \rangle$. We get

$$\begin{aligned} \langle \Omega | \phi(x) \phi(y) | \Omega \rangle &= \langle \Omega | \phi(x) | \Omega \rangle \langle \Omega | \phi(y) | \Omega \rangle \\ &+ \int \frac{d^3 \mathbf{p}}{(2\pi)^3 (2\omega_{\mathbf{p}})} \langle \Omega | \phi(x) | \mu_{\mathbf{p}}^2 \rangle \langle \mu_{\mathbf{p}}^2 | \phi(y) | \Omega \rangle \\ &+ \int_{M_0^2}^{\infty} d(M^2) \int \frac{d^3 \mathbf{p}}{(2\pi)^3 (2\omega_{\mathbf{p}}(M^2))} \langle \Omega | \phi(x) | M_{\mathbf{p}}^2 \rangle \langle M_{\mathbf{p}}^2 | \phi(y) | \Omega \rangle. \end{aligned}$$

Part of our definition of an interpolating field required that $\langle \Omega | \phi(x) | \Omega \rangle = 0$, so the first term vanishes. For the rest, I encourage you to verify that

$$\langle \Omega | \phi(x) | M_{\mathbf{p}}^2 \rangle = \langle \Omega | \phi(0) | M_{\mathbf{0}}^2 \rangle e^{-ip \cdot x},$$

where the time component of the p in the final expression is $\omega_{\mathbf{p}}(M^2)$, and that this is true both when $M^2 = \mu^2$ and when it's one of the values from the continuous part of the spectrum. (It will be helpful to use the fact that $\phi(x) = e^{iP \cdot x} \phi(0) e^{-iP \cdot x}$, and that if B_p is the unitary operator corresponding to the Lorentz boost that takes p to $(p^2, 0, 0, 0)$, then since 0 is fixed by this boost, $B_p^{-1} \phi(0) B_p = \phi(0)$.)

We therefore have

$$\begin{aligned} \langle \Omega | \phi(x) \phi(y) | \Omega \rangle &= \int \frac{d^3 \mathbf{p}}{(2\pi)^3 (2\omega_{\mathbf{p}})} |\langle \Omega | \phi(0) | \mu_{\mathbf{0}}^2 \rangle|^2 e^{-ip \cdot (x-y)} \\ &+ \int_{M_0^2}^{\infty} d(M^2) \int \frac{d^3 \mathbf{p}}{(2\pi)^3 (2\omega_{\mathbf{p}}(M^2))} |\langle \Omega | \phi(0) | M_{\mathbf{0}}^2 \rangle|^2 e^{-ip \cdot (x-y)}. \end{aligned}$$

If we now assume that $x^0 > y^0$, since the factor involving ϕ is now independent of \mathbf{p} , the \mathbf{p} integral is just equal to the Feynman propagator, except with M^2 in place of m^2 . Therefore, if we write

$$\rho(M^2) = |\langle \Omega | \phi(0) | M_{\mathbf{0}}^2 \rangle|^2 + |\langle \Omega | \phi(0) | \mu_{\mathbf{0}}^2 \rangle|^2 \delta(M^2 - \mu^2),$$

where the first term is taken to be 0 when $M^2 < M_0^2$, then

$$\langle \Omega | T[\phi(x) \phi(y)] | \Omega \rangle = \int_0^{\infty} d(M^2) \rho(M^2) \Delta_F^M(x-y).$$

(We've verified this when $x^0 > y^0$; I encourage you to make sure everything still works in the other case.)

This is called the **Källén–Lehmann spectral representation** of the time-ordered two-point function, and ρ is called the **spectral density**.

6.2 Interpreting the Spectral Representation

It's useful to think of the Källén–Lehmann representation as a sort of replacement for the simpler relation $\langle 0 | T[\phi(x) \phi(y)] | 0 \rangle = \Delta_F^m(x-y)$ that we had in the free theory. As we discussed a bit in the last article, while in the free theory $\phi(x) | 0 \rangle$ lives completely in the one-particle subspace, in

the interacting theory there is some “leakage” into the multi-particle subspace. When analyzing the two-point function, then, we don’t only get the propagator $\Delta_F^\mu(x-y)$ that we might expect; we also get extra contributions from the other components of $\phi(x)|\Omega\rangle$.

It’s useful to see what this tells us about the Fourier-transformed two-point function $\tilde{G}^{(2)}$. First of all, the coefficient in front of the delta function portion of the spectral density ρ is $|\langle\Omega|\phi(0)|\mu_0^2\rangle|^2$. But recall that we defined our field normalization Z by setting $\langle p|\phi(0)|\Omega\rangle = Z^{1/2}$. Therefore, the coefficient on the delta function contribution to ρ is just Z . (This is the reason for the one-half power in the original definition.) From our definition above,

$$\tilde{G}^{(2)}(p, -q) = \delta(p - q) \int_0^\infty dM^2 \rho(M^2) \frac{i}{p^2 - M^2 + i\epsilon}.$$

We can split off the contribution to ρ from the one-particle states and write

$$\tilde{G}^{(2)}(p, -q) = \delta(p - q) \left[\frac{iZ}{p^2 - \mu^2 + i\epsilon} + \int_{M_0^2}^\infty dM^2 \rho(M^2) \frac{i}{p^2 - M^2 + i\epsilon} \right].$$

This gives us a nice way to extract information about μ and Z from $\tilde{G}^{(2)}$: the factor multiplying the delta function in $\tilde{G}^{(2)}(p, -q)$ has a pole at $p^2 = \mu^2$, and Z is the residue at that pole. You could compute μ and Z to any degree in perturbation theory by taking the reciprocal of the corresponding approximation to the factor multiplying the delta function in $\tilde{G}^{(2)}(p, -q)$, finding where this reciprocal is equal to zero to identify μ , and then taking the derivative at $p^2 = \mu^2$ to identify Z . (We’ll have a somewhat more efficient method when we tackle this computation in the next article.)

There might also, conceivably, be other isolated eigenvalues of P^2 between μ^2 and M_0^2 . Physicists refer to the corresponding eigenstates as **bound states** of our theory, and they are essentially other one-particle states. For simplicity we assumed that there were no such eigenvalues, but our analysis can accommodate them quite straightforwardly: each particle state which has nonzero overlap with $\phi(x)|\Omega\rangle$ will add an extra delta function onto the spectral density ρ , and therefore an extra pole to the $\tilde{G}^{(2)}$. This gives a very nice way to think about the relationship between particles and their interpolating fields: particles for which a field operator A is an interpolating field correspond to poles in the two-point function for A .

7 Everything We Just Did Is Garbage

In some ways our story is quite nice: we started with a quantity involving a quantum-mechanical system that we had no hope of solving precisely and managed to find an algorithm for approximating this quantity to any desired degree in perturbation theory in terms of a sum of integrals associated to Feynman diagrams.

But there is one very unfortunate fact about the integrals that arise from this process: they have a tendency to diverge. We already caught a glimpse of this in our discussion of vacuum bubbles, but the problem is actually much more pervasive than that. Recall that the example we discussed when deriving the momentum-space Feynman rules contained the following integral:

$$\lim_{\epsilon \rightarrow 0^+} \int \frac{d^4k}{(2\pi)^4} \frac{i}{k^2 - m^2 + i\epsilon} \frac{i}{(p_1 + p_2 - k)^2 - m^2 + i\epsilon}.$$

For large k , that denominator grows like k^4 , so this integral definitely diverges. While we were able to find a way to eliminate the “infinities” that corresponded to our vacuum bubbles,

it's much harder to do that for this diagram; there's no other diagram around that obviously cancels the contribution from this one. What's worse, the sort of divergence that shows up here is typical of the integrals that arise from Feynman diagrams containing at least one loop. In most interacting quantum field theories not only will the story we just told produce divergent integrals, it will do so for *infinitely many* of the diagrams we are asked to draw!

This is, needless to say, quite problematic, and it raises two obvious questions: why did this happen, and what are we going to do about it? I want to close this article by saying a few words about both of these questions, although a complete answer to the second one will be the main topic of the next article in this series.

Let's start with the question of why. There are three different mathematical problems that I'd like to mention now:

- The first and most serious problem is that the operators we manipulated to arrive at our Feynman rules do not actually exist. As we alluded to when we first defined the Hamiltonian, the field operators $\phi(x)$ are operator-valued distributions, and so products of them (not to mention fourth powers!) should not be expected to be well-defined in general. Recall once again that we saw a manifestation of this even for the free field in the first article, and even before removing the constraint that our fields were confined to a box. Since our strategy for dealing with this was to just press ahead and multiply distributions as though they were functions, we shouldn't be surprised that most of the resulting computations produced nonsense answers.
- The previous issue affects the computations we performed *after* reducing everything to the free theory, but there is another, somewhat subtler problem that affects the reduction itself. We started that story by supposing that there was a single Hilbert space on which we had defined field operators $\phi(\mathbf{x})$, along with two different Hamiltonians H_0 and H which we used to translate the field operators in time in two different ways. We then posited the existence of two different vacuum states, $|0\rangle$ for H_0 and $|\Omega\rangle$ for H , and asked for the overlap $\langle 0|\Omega\rangle$ to be nonzero.

Unfortunately, even if we could manage to formally define H in a satisfactory way (which, again, we can't) the setup just described is actually known to be impossible due to a result called *Haag's Theorem*. If it were possible to realize the free and interacting theories in Hilbert spaces separately, then we could of course combine them into a single Hilbert space simply by taking the direct sum, but that would imply that $\langle 0|\Omega\rangle = 0$, which would break the story we're trying to tell. In this context, Haag's Theorem amounts to the statement that this orthogonality of the two vacuum states is actually an inevitable consequence of the rest of our assumptions. For more on this, I recommend the discussion in Section 10.5 of Anthony Duncan's *The Conceptual Framework of Quantum Field Theory*.

- Finally, there is the issue of the convergence of the perturbation series itself. When we wrote down the Dyson series for $V(t, t')$ and the associated perturbation series representation of the n -point function, we implicitly assumed that everything converged. If the interaction part H_I of the Hamiltonian were a bounded operator on an honest Hilbert space, it is possible to show that the Dyson series indeed converges in the operator norm topology. But even if H_I were a well-defined operator it certainly wouldn't be bounded.

Even if we could somehow fix the other two issues listed here, there doesn't seem to be a good reason to expect the perturbation series to actually converge, and in fact there are some good physical reasons to believe it shouldn't. (There is a nice, if short, discussion of

this on p. 253 of Folland's book, which includes a reference to a paper of Dyson's making the original argument.)

I don't know of any actual theorems in this area, but I believe the best that anyone currently hopes for is that the perturbation series is an asymptotic series but not a convergent one. This still leaves open the possibility that (once we have found some way to extract a finite number from them) the first few terms of the series are a decent approximation of the relevant physical quantities, and this is how all actual numerical computations in this area go — rather than worry about convergence, physicists just compute as far out in the perturbation series as is computationally feasible and hope for the best.

These mathematical issues do something to explain why we ended up with divergent integrals. But what are we supposed to do about it? The problem of the non-convergence of the perturbation series has, to my knowledge, no more satisfying resolution than “take the first few terms and hope,” but if we restrict ourselves to the problem of assigning a meaning to each term in the series *separately* then there is much more to say.

As we said when discussing vacuum bubbles, our strategy will be to take these divergent integrals and regard them as limits of some finite quantities as some parameter (the *cutoff*) goes to infinity. The process of producing these finite quantities is referred to as *regularizing* the integral, and there are many different regularization schemes one can use, each with its own set of computational tradeoffs.

Of course, if this were all we were doing, it would be no help at all: the values of the regularized integrals would depend on the cutoff, and if the regularized integrals are to be regarded as approximations of the original divergent integral then we should also expect this value to blow up as we take the cutoff to infinity. This would be like expecting the sum $\sum_{k=1}^{\infty} k$ to start producing a usable result just because we chose to write it as $\lim_{n \rightarrow \infty} \frac{n(n+1)}{2}$.

Instead, we will be saved by a new observation about these computations. Our original Lagrangian depends on two parameters: the mass m and the so-called “coupling constant” λ . The crucial thing to realize here is that these parameters are not actually physically meaningful quantities. We discussed at some length in the previous article that m is under no obligation to line up with the mass of the particle, which we called μ , and in an analogous way λ is also not directly measurable in terms of an actual physical process.

What this means is that if we perform our limit in the way you were probably imagining — holding m and λ constant, computing the values of our regularized integrals as functions of m , λ , and the cutoff, and then sending the cutoff to infinity — we are in an important sense *asking the wrong question*. Why insist on a fixed value for m and λ when I have no idea what they have to do with the physical situation I'm modeling? It would make much more sense to instead insist on fixed values of quantities that actually have physical meaning, like, for example, the mass μ of the actual particle.

If we take the limit in this way instead, we will find that m and λ will, in general, depend on the cutoff. But the hope is that our regularized Feynman integrals, *when expressed as functions of our new, more physically meaningful parameters*, will now converge to something finite and sensible as the cutoff goes to infinity. When everything goes right, this is exactly what happens.

While I hope that brief discussion was helpful, there really is no substitute for seeing this process carried out in an explicit computation involving a specific regularization scheme. This is exactly the task we'll take up in the next article in this series.