Math 273B: Topics in Mathematical Physics

Lecturer: Professor Sourav Chatterjee

Notes by: Andrew Lin

Winter 2025

Introduction

This is the continuation of last quarter's Math 273A course. We'll begin with a review of what we did in the fall (but notes are on Canvas if we want more details, and they'll be updated each week with new material). And if you're reading this set of notes, it might be advised to read the Math 273A notes first!

1 January 6, 2025

Our goal is to rigorously build quantum field theories, and in general the setting of such a **quantum system** is that we have a Hilbert space $\mathcal{H}$ along with a strongly continuous one-parameter group of unitary operators $(U(t))_{t\in\mathbb{R}}$ (we'll explain what this means shortly).

Definition 1

A **state** of the system is a ray of the Hilbert space $\mathcal{H}$ – that is, a set of the form $\{\alpha\psi : \alpha \in \mathbb{C}\setminus\{0\}\}$ for some nonzero $\psi \in \mathcal{H}$.

In other words, two vectors that are nonzero scalar multiples of each other represent the same state. The way this group of unitary operators acts on our states is that if our system is at state ψ at time s, then it is at state $U(t)\psi$ at time $s+t$ (and we'll really think of states as just elements of $\mathcal{H}$ when it doesn't cause confusion). Recall that a **unitary operator** is a linear function $U : \mathcal{H} \to \mathcal{H}$ such that $UU^* = U^*U = I$ (where $U^* : \mathcal{H} \to \mathcal{H}$ is the **adjoint** operator of U, which is the map satisfying $\langle Ux, y\rangle = \langle x, U^*y\rangle$ for all $x, y \in \mathcal{H}$; any linear map has such an adjoint). So in particular we are saying that our family $(U(t))_{t\in\mathbb{R}}$ satisfies the following:

- For all t, the operator $U(t)$ is unitary,
- $U(0) = I$ (though this also follows from the other conditions),
- $U(s+t) = U(s)U(t)$,
- (Strong continuity condition) For all $x \in \mathcal{H}$, the map $t \mapsto U(t)x$ is continuous (from $\mathbb{R}$ into $\mathcal{H}$).

Note that we're not discussing relativistic systems at the moment (if we did, we'd also need to worry about Lorentz invariance, but our setup will suffice for now). So knowing our state at any time tells us the state at all past or future times via this evolution.

One of the foundational theorems we proved last quarter was Stone's theorem, and to state this we need the notion of a **self-adjoint** operator. Recall that a **densely-defined** linear operator on $\mathcal{H}$ is a (not necessarily continuous) linear

map $T : \mathcal{D} \to \mathcal{H}$, where $\mathcal{D}$ is a dense subspace of $\mathcal{H}$ (called the **domain** of T). The adjoint T^* of such an operator is now defined by first specifying the domain of T^* to be

$$\mathcal{D}(T^*) = \{y \in \mathcal{H} : \exists z \text{ such that } \langle Tx, y\rangle = \langle x, z\rangle \ \forall x \in \mathcal{H}\},$$

and for all such y in this domain we let $T^*y = z$ (it's easy to see that such a z is unique if it exists, and also that $\mathcal{D}(T^*)$ is indeed a subspace of $\mathcal{H}$). In general it can be difficult to relate the domains of T and T^*, but in the case where T is **symmetric** (also **Hermitian**) – meaning that $\langle Tx, y\rangle = \langle x, Ty\rangle$ for all $x, y \in \mathcal{D}(T)$ – we can say more. Specifically, we have $\mathcal{D}(T) \subseteq \mathcal{D}(T^*)$, so T^* will also be densely defined if T is. And we say that T is **self-adjoint** if $\mathcal{D}(T) = \mathcal{D}(T^*)$.

Example 2

Consider the second-derivative operator $\frac{d^2}{dx^2}$ on $\mathbb{R}$. This acts on a dense subset of L^2, the set of smooth compactly supported functions. By integration by parts, we can see that this operator is indeed symmetric (since $\langle x, y'''\rangle = \langle x'', y\rangle$). But this is not self-adjoint, since the domain is too small – T^* will be defined on a bigger set. So we can keep getting bigger and bigger domains by repeatedly taking adjoints, and the limiting domain is the set of $\hat{\psi}$ such that $\int p^4|\hat{\psi}|^2 dp < \infty$. So self-adjointness gives us the "maximal extension."

We'll restate Stone's theorem here for its use this quarter:

Theorem 3 (Stone)

Let $\mathcal{H}$ be a Hilbert space, and let $(U(t))_{t\in\mathbb{R}}$ be a strongly continuous one-parameter group of unitary operators on $\mathcal{H}$. Let

$$\mathcal{D} = \left\{x \in \mathcal{H} : \lim_{t\to 0} \frac{U(t)x - x}{-it} \text{ exists}\right\}.$$

For all $x \in \mathcal{D}$, let Hx be this limit. Then $\mathcal{D}$ is dense, and H is a (densely-defined) self-adjoint operator with domain $\mathcal{D}$ which we call the **Hamiltonian**. Conversely, let H be any densely-defined self-adjoint operator with domain $\mathcal{D}$. Then there exists a unique one-parameter strongly-continuous group of unitary operators $(U(t))_{t\in\mathbb{R}}$, such that $\mathcal{D}$ and H arise in the above way from that group.

The relation we should keep in mind here is "$U(t) = e^{-itH}$," and this comes from the differential equation (called the **Schrodinger equation**) $\frac{d}{dt}U(t)x = -iHU(t)x$ (the right-hand side makes sense because $U(t)$ will always map into the domain of H). And if H is bounded and defined on the full Hilbert space, then $U(t) = e^{-itH}$ actually makes sense as an infinite sum.

In general, proving self-adjointness of an operator H can be very difficult to do (and there are other issues too, such as understanding the spectrum of H beyond the fact that it's contained in the real line). The way we intend to proceed is to instead use probability theory via what's called the "Wick rotation heuristic." We call a family $(P(t))_{t\geq 0}$ of linear maps from $\mathcal{H}$ into $\mathcal{H}$ a **semigroup** if $P(0) = I$ and $P(s + t) = P(s)P(t)$ for all $s, t \geq 0$, and we say it is **strongly continuous** if for all $x \in \mathcal{H}$, the map $t \mapsto P(t)x$ is continuous. (So far, this is the same definition as before but only with nonnegative s, t.) But we now want to consider symmetric, rather than unitary, operators:

Theorem 4 (Hille–Yosida theorem, one direction)

Let $(P(t))_{t\geq 0}$ be a strongly continuous semigroup of symmetric operators on a Hilbert space $\mathcal{H}$. Then

$$\mathcal{D} = \left\{ x \in \mathcal{H} : \lim_{t \downarrow 0} \frac{U(t)x - x}{-t} \text{ exists} \right\}$$

is dense in $\mathcal{H}$, and if we let Hx be this limit for all $x \in \mathcal{D}$, then H is self-adjoint with domain $\mathcal{D}$.

So we get a second way of arriving at our operator H, and this type the differential equation we satisfy is $\frac{d}{dt}P(t)x = -HP(t)x$ (so that we write $P(t) = e^{-tH}$ instead of $U(t) = e^{-itH}$). This is the whole point of Wick rotation – we'll do an example in detail soon, but the point is that we can define P via a stochastic process so that everything is mathematically rigorous. The converse direction of Stone's theorem then tells us that $U(t)$ has to also be well-defined. (In physics, we write down the path integral for $U(t)$ and then replace t with $-it$ everywhere except in the arguments of functions so that e^{-itH} becomes e^{-tH}, but what we're doing is a rigorous version of that.)

Example 5

The system we'll focus on first is a single particle on $\mathbb{R}$ moving in a potential $V : \mathbb{R} \to \mathbb{R}$. Setting Planck's constant to $\hbar = 1$ for simplicity, the Schrodinger equation reads

$$i\frac{\partial \psi}{\partial t} = -\frac{1}{2}\frac{\partial^2 \psi}{\partial x^2} + V(x)\psi(t,x).$$

To do things rigorously, we need to write down what the relevant spaces are: we can take $H = L^2(\mathbb{R})$, and we can define this evolution for smooth compactly supported ψ. But then we need to turn it into a self-adjoint operator so that we can use Stone's theorem, and that's what we'll do here. In particular, this doesn't work for arbitrary V, and we'll do it for the case where V is **continuous and bounded below** (otherwise the particle might escape to infinity).

Recall from last quarter that when we tried writing down the oscillatory path integral, we could only solve things because we had a Gaussian integral and still had to do some work with contour integration. With a general V or even for something like a quartic potential, the problem stops looking solvable. But it'll turn out that moving to the probabilistic interpretation (that is, the Euclidean system), we can write everything for $P(t)$ down explicitly in terms of Brownian motion (which is a rigorously defined mathematical object which we'll define) using the Feynman-Kac formula. That'll tell us that there is indeed a subspace of L^2 so that the right-hand side above, $H\psi(x)$, can be extended to a subspace so that we have global solutions for all time.

We won't dwell on this example too much, though, since this isn't a relativistic system. Specifically, we have $\psi(t,x)$ evolving according to this equation, but if we do a Lorentz transformation of spacetime, then physics should remain invariant. But the wavefunction's transformed evolution no longer evolves according to the Schrodinger equation, and thus this is somehow "not physical." So immediately after this equation was introduced, there was a lot of activity making this relativistic; currently there is no system with finitely many particles that we know of that obeys all of the restrictions imposed by relativity. So that's why quantum field theory is the direction we'll be going in over the next few weeks!

2 January 8, 2025

Last time, we were considering the case of a single particle on the real line in a potential $V : \mathbb{R} \to \mathbb{R}$. The path integral prescription of the evolution is the following: we should have

$$\psi_t(x) = \frac{1}{Z(t)} \int_{\text{paths } \gamma:\gamma(t)=x} \psi_0(\gamma(0)) e^{iS(\gamma)} \mathcal{D}\gamma$$

where $Z(t)$ is some appropriately chosen normalizing constant, $S(\gamma) = \int_0^t \left(\frac{1}{2}\dot{\gamma}(s)^2 - V(\gamma(s))\right) ds$ is the action whose critical points are the classical trajectories (since Newton's equation of motion says that $\ddot{\gamma}(s) = -V'(\gamma(s))$ for a particle of mass 1), and $\mathcal{D}\gamma$ is some "Lebesgue measure on the space of paths" (which doesn't make mathematical sense at the moment). Note that we usually choose $Z(t)$ so that the evolution preserves the L^2 norm, since the normalization doesn't actually change what state we're in.

Remark 6. *It's possible to construct Lebesgue measure-type objects on path spaces, but they tend to be very ill-behaved because we can't apply the usual tools of probability theory. Instead, we'll see other ways to make sense of these things over the next few weeks.*

We'll mostly do heuristics today, generally describing how we get the Schrodinger equation out of the path integral and what Hamiltonian arises out of that. We'll then see how to heuristically do a Wick rotation to get to a non-oscillatory integral.

First of all, note that this is not the usual way that path integrals are written. Reversing time on the interval $[0, t]$, we may also equivalently write

$$\psi_t(x) = \frac{1}{Z(t)} \int_{\gamma:\gamma(0)=x} \psi_0(\gamma(t)) e^{iS(\gamma)} \mathcal{D}\gamma$$

(the action doesn't change because we square the derivative in our expression for S). But intuitively the first description makes more sense, since it's like a probability transition (that is, we look at all possible starting points and find the probability of going from that starting point to our final location). The choice of Z that will turn out to give us a unitary evolution is

$$Z(t) = \int_{\gamma:\gamma(0)=0} e^{iS_0(\gamma)} \mathcal{D}\gamma,$$

where $S_0(\gamma) = \frac{1}{2}\int_0^t \dot{\gamma}(s)^2 ds$ is the "kinetic part" of our evolution. (Notice that getting a unitary evolution instead of a "probability" one means we have to pick S_0 instead of S in the exponent for normalization, and we'll soon see why.)

Example 7

We'll show now (heuristically) that the map $U(t)$ mapping ψ_0 to ψ_t is indeed a one-parameter group of unitary evolutions.

First of all, to see that the normalizing constants satisfy $Z(t+s) = Z(t)Z(s)$ (this will turn out to be important after Wick rotation), we first decompose a path $\gamma : [0, t+s] \to \mathbb{R}$ into two parts, namely $\alpha : [0, t] \to \mathbb{R}$ concatenated with $\beta : [t, t+s] \to \mathbb{R}$, such that $\alpha(0) = 0$ and $\alpha(t) = \beta(t)$. If we then define $\theta : [0, s] \to \mathbb{R}$ so that $\theta(u) = \beta(t+u) - \alpha(t)$ for all $u \in [0, s]$ (that is, translating the path in space and time to start at the origin), we have $\dot{\theta}(u) = \dot{\gamma}(t+u)$ and thus we may write

$$Z(t+s) = \int_{\alpha:\alpha(0)=0} \int_{\theta:\theta(0)=0} \exp\left(\frac{i}{2}\int_0^t \dot{\alpha}(u)^2 du + \frac{i}{2}\int_0^s \dot{\theta}(u)^2 du\right) \mathcal{D}\alpha\mathcal{D}\theta,$$

which is the product of two integrals (over α and β) which are then $Z(t)$ and $Z(s)$, as desired.

A similar thing also works for the numerator (but we have to be a bit more careful): defining

$$Q(t)\psi(x) = \int_{\gamma:\gamma(0)=x} \psi(\gamma(t))e^{iS(\gamma)}\mathcal{D}\gamma,$$

we claim that the operators $Q(t)$ also satisfy $Q(t+s) = Q(t)Q(s)$. Now writing γ as a concatenation of α and δ, where $\alpha(u) = \gamma(u)$ for $0 \le u \le t$ and $\delta(u) = \gamma(t+u)$ (we don't want to subtract off $\alpha(t)$ this time because the potential is not invariant under translations), we must have $\alpha(0) = x$ and $\delta(0) = \gamma(t)$. We thus have

$$Q(t+s)\psi(x) = \int_{\alpha:\alpha(0)=x} \int_{\delta:\delta(0)=\alpha(t)} \psi(\delta(s))e^{iS(\alpha)+iS(\delta)}\mathcal{D}\alpha\mathcal{D}\delta$$

and now the two integrals can't just be interchanged anymore, but the action does split up additively. Now we can bring the $e^{iS(\alpha)}$ factor outside, leaving us with

$$\int_{\alpha:\alpha(0)=x} e^{iS(\alpha)} \left(\int_{\delta:\delta(0)=\alpha(t)} \psi(\delta(s))e^{iS(\delta)}\mathcal{D}\delta \right) \mathcal{D}\alpha,$$

and now the inner parenthetical term is $Q(s)\psi(\alpha(t))$ (since our paths are starting at $\alpha(t)$ so that we have

$$Q(t+s)\psi(x) = \int_{\alpha:\alpha(0)=x} (Q(s)\psi)(\alpha(t))e^{iS(\alpha)}\mathcal{D}\alpha = (Q(t)(Q(s)\psi))(x),$$

so that indeed $Q(t+s) = Q(t)Q(s)$, as desired.

We can now do the full calculation: the operators we are considering are

$$U(t)\psi(x) = \frac{1}{Z(t)}Q(t)\psi(x),$$

and we wish to show that $U(t+s) = U(t)U(s)$. Indeed,

$$\begin{aligned} U(t+s)\psi(x) &= \frac{1}{Z(t+s)}Q(t+s)\psi(x) \\ &= \frac{1}{Z(t)Z(s)}(Q(t)(Q(s)\psi))(x), \end{aligned}$$

but because $Q(t)$ is a linear operator we can move the $Z(s)$ constant inside and get

$$\frac{1}{Z(t)}Q(t)\left(\frac{1}{Z(s)}Q(s)\psi\right)(x) = U(t)(U(s)\psi)(x),$$

showing the group property. From here, we'll compute (heuristically) the Hamiltonian, which will help us complete the rest of the proof of unitarity – recall that we should have $H\psi(x) = \lim_{t\to 0} \frac{U(t)\psi(x)-\psi(x)}{-it}$. This next part will be rather handwavy: when t is small, we expect that "most paths γ remain close to the starting point $\gamma(0)$ for all $s \in [0,t]$," and the "far paths" have a large enough action that the oscillations will cancel out. (No one really knows how to make this rigorous, since no one can simulate Feynman path integrals in a good way like we can simulate solutions to SDEs.) We'll say that for a typical path $\gamma : [0,t] \to \mathbb{R}$ with $\gamma(0) = x$, we expect

$$\int_0^t V(\gamma(s))ds = V(x)t + o(t)$$

and (we'll see later why going up to second order is the right amount)

$$\psi(\gamma(t)) = \psi(x) + (\gamma(t)-x)\psi'(x) + \frac{1}{2}(\gamma(t)-x)^2\psi''(x) + \text{higher order terms which are } o(t).$$

Plugging these in, we thus have

$$\begin{aligned} Q(t)\psi(x) &= \int_{\gamma:\gamma(0)=x} \psi(\gamma(t)) \exp\left(iS_0(\gamma) - i\int_0^t V(\gamma(s))ds\right)\mathcal{D}(\gamma) \\ &= \int_{\gamma:\gamma(0)=x} \left(\psi(x) + (\gamma(t)-x)\psi'(x) + \frac{1}{2}(\gamma(t)-x)^2\psi''(x) + o(t)\right)\exp\left(iS_0(\gamma) - itV(x) + o(t)\right)\mathcal{D}(\gamma) \end{aligned}$$

and now we can subtract off x and shift γ to zero; this doesn't affect $S_0(\gamma)$ because that only depends on $\dot{\gamma}$, and we end up with

$$Q(t)\psi(x) = \exp\left(-itV(x) + o(t)\right)\int_{\gamma:\gamma(0)=0} (\psi(x) + \gamma(t)\psi'(x) + \frac{1}{2}\gamma(t)^2\psi''(x) + o(t))e^{iS_0(\gamma)}\mathcal{D}\gamma.$$

A similar calculation with the normalizing factor shows that $Z(t) = \int_{\gamma:\gamma(0)=0} e^{iS_0(\gamma)}\mathcal{D}\gamma$, so that when we divide these two expressions we get

$$U(t)\psi(x) = e^{-itV(x)+o(t)}\left(\psi(x) + \psi'(x)\frac{\int \gamma(t)e^{iS_0(\gamma)}\mathcal{D}\gamma}{\int e^{iS_0(\gamma)}\mathcal{D}\gamma} + \frac{1}{2}\psi''(x)\frac{\int \gamma(t)^2 e^{iS_0(\gamma)}\mathcal{D}\gamma}{\int e^{iS_0(\gamma)}\mathcal{D}\gamma} + \cdots\right),$$

and each of these ratios $\frac{\int \gamma(t)^k e^{iS_0(\gamma)}\mathcal{D}\gamma}{\int e^{iS_0(\gamma)}\mathcal{D}\gamma}$ can indeed be exactly evaluated by the discretization and Gaussian integral methods we discussed last quarter – it turns out to equal 0 for $k=1$ and it for $k=2$. (The higher powers then turn out to exactly be higher powers of t, so we can indeed ignore them.) Therefore we actually have

$$\begin{aligned} U(t)\psi(x) &= e^{-itV(x)+o(t)}\left(\psi(x) + \frac{1}{2}it\psi''(x) + o(t)\right) \\ &= (1 - itV(x) + o(t))\left(\psi(x) + \frac{1}{2}it\psi''(x) + o(t)\right) \\ &= \psi(x) + \frac{1}{2}it\psi''(x) - itV(x)\psi(x) + o(t), \end{aligned}$$

and therefore plugging back into the definition of the Hamiltonian we get $\boxed{H(x) = -\frac{1}{2}\psi''(x) + V(x)}$, and we can verify that $\frac{\partial}{\partial t}U(t)\psi = -iHU(t)\psi$ is indeed the Schrodinger equation we promised. So choosing $Z(t)$ helped us out a lot here in making the exact calculations work out, and it'll help us out further later.

The useful fact now is that H is actually symmetric – indeed, by integration by parts and the fact that V is real-valued, we have

$$\langle H\psi, \phi\rangle = \langle \psi, H\phi\rangle,$$

where we're taking the inner product of complex-valued functions $\langle f, g\rangle = \int_{-\infty}^{\infty} f(x)\overline{g(x)}dx$. So together with the Schrodinger equation $\frac{d}{dt}U(t)\psi = -iHU(t)\psi$, we find that

$$\begin{aligned} \frac{d}{dt}\langle U(t)\psi, U(t)\phi\rangle &= \left\langle \frac{d}{dt}U(t)\psi, U(t)\phi\right\rangle + \left\langle U(t)\psi, \frac{d}{dt}U(t)\phi\right\rangle \\ &= \langle -iHU(t)\psi, U(t)\phi\rangle + \langle U(t)\psi, -iHU(t)\phi\rangle \\ &= -i\langle HU(t)\psi, U(t)\phi\rangle + i\langle U(t)\psi, HU(t)\phi\rangle \\ &= -i\langle HU(t)\psi, U(t)\phi\rangle + i\langle HU(t)\psi, U(t)\phi\rangle \\ &= 0 \end{aligned}$$

by conjugate linearity and then symmetry. So $U(t)$ is indeed an isometry; it takes more work to show from here that U is surjective, but we'll just show that the Hamiltonian is self-adjoint and that will do the job for us.

3 January 10, 2025

We'll start making things rigorous today. Recall that we've been considering a single particle in a potential and want to construct solutions to the Schrodinger equation $i\frac{\partial\psi}{\partial t} = -\frac{1}{2}\frac{\partial^2\psi}{\partial x^2} + V(x)\psi(t,x)$. Specifically, we want to define the operator on the right-hand side, $H\psi = -\frac{1}{2}\psi'' + V\psi$, on a large enough domain so that we actually get self-adjointness. One way to do this (by Hille–Yosida) is to get H instead as the generator of a strongly continuous semigroup $(P_t)_{t\geq 0}$ of **symmetric** operators on $L^2(\mathbb{R})$, specifically $P(t) = e^{-tH}$. We'll be using probability theory to rigorously construct such a semigroup.

Since $U(t) = e^{-itH}$, the idea will be to take the path integral heuristic and replace t with $-it$ everywhere it occurs in a computation (but not when it's an argument of a function). So if we fix some time $t > 0$, the action of our trajectory will be

$$S(\gamma) = \int_0^t \left(\frac{1}{2}\dot\gamma(s)^2 - V(\gamma(s))\right) ds,$$

and then (remembering that S_0 denotes only the $\frac{1}{2}\dot\gamma^2$ part of the integral)

$$U(t)\psi(x) = \frac{\int_{\gamma:\gamma(0)=x} \psi(\gamma(t)) e^{iS(\gamma)}\mathcal{D}\gamma}{\int_{\gamma:\gamma(0)=0} e^{iS_0(\gamma)}\mathcal{D}\gamma}.$$

We'll now see how this gets transformed under "Wick rotation" $s \mapsto -is$. Since "$ds = (s+\varepsilon) - s$ for ε infinitesimally small," we should replace $ds \mapsto -i\,ds$, and similarly $\dot\gamma(s) = \lim_{\varepsilon\to 0}\frac{\gamma(s+\varepsilon)-\gamma(s)}{\varepsilon}$ will become $\lim_{\varepsilon\to 0}\frac{\gamma(s+\varepsilon)-\gamma(s)}{-i\varepsilon} = +i\dot\gamma(s)$. Therefore our action transforms as (notice that the s inside γ doesn't get replaced because it's the argument of a function)

$$S(\gamma) \mapsto \int_0^t \left(-\frac{1}{2}\dot\gamma(s)^2 - V(\gamma(s))\right)(-i\,ds) = i\int_0^t \left(\frac{1}{2}\dot\gamma(s)^2 + V(\gamma(s))\right) ds.$$

Similarly $S_0(\gamma)$ will become $i\int_0^t \frac{1}{2}\dot\gamma(s)^2 ds$, and thus plugging everything back in we see that

$$U(t)\psi(x) \mapsto \frac{\int_{\gamma:\gamma(0)=x} \psi(\gamma(t)) e^{-S_E(\gamma)}\mathcal{D}\gamma}{\int_{\gamma:\gamma(0)=0} e^{-S_0(\gamma)}\mathcal{D}\gamma},$$

where S_E is now the **Euclidean action**

$$S_E(\gamma) = \int_0^t \left(\frac{1}{2}\dot\gamma(s)^2 + V(\gamma(s))\right) ds.$$

Of course, this expression doesn't really make any more sense than the previous version, but the prescription tells us what to aim for in the probabilistic setting – we should notice that the is are all gone so that there's no oscillatory integrals, and the $-V$ is now replaced with a $+V$. This is what we'll be trying to make rigorous.

Example 8

Towards rigor, we'll do the discretization strategy we did before and try to take a limit.

Instead of the full time-interval $[0,t]$, we'll consider discrete time points $0, \frac{t}{n}, \frac{2t}{n}, \cdots$ and take $n\to\infty$; let $x_j = \gamma(\frac{jt}{n})$ (for $0 \leq j \leq n$) be the discretization of the path γ. Our integral $\mathcal{D}\gamma$ over paths is then replaced with the Lebesgue integral over $x_1, x_2, \cdots, x_n$ (since $\gamma(0) = x_0$ is fixed but all other points are free). We'll then replace

$$\dot\gamma\left(\frac{jt}{n}\right) \mapsto \frac{\gamma\left(\frac{(j+1)t}{n}\right) - \gamma\left(\frac{jt}{n}\right)}{\frac{t}{n}} = \frac{n}{t}(x_{j+1} - x_j),$$

so that the discretized Euclidean action reads (the different summations are just to keep expressions simpler later)

$$\begin{aligned} S_E(\gamma) &\mapsto \frac{t}{n}\sum_{j=0}^{n-1}\left(\frac{1}{2}\left(\frac{n}{t}(x_{j+1}-x_j)\right)^2 + \frac{t}{n}\sum_{j=1}^{n} V(x_j)\right) \\ &= \frac{n}{2t}\sum_{j=0}^{n-1}(x_{j+1}-x_j)^2 + \frac{t}{n}\sum_{j=1}^{n} V(x_j). \end{aligned}$$

Very similarly we can discretize the denominator (note that now x_0 is fixed at x instead of 0), so that the approximation of P is

$$P_n(t)\psi(x) = \frac{\int\cdots\int \psi(x_n)\exp\left(-\frac{n}{2t}\sum_{j=0}^{n-1}(x_{j+1}-x_j)^2 - \frac{t}{n}\sum_{j=1}^{n} V(x_j)dx_1\cdots dx_n\right)}{\int\cdots\int \exp\left(-\frac{n}{2t}\sum_{j=0}^{n-1}(z_{j+1}-z_j)^2 dz_1\cdots dz_n\right)}$$

where the only difference between xs and zs is that $x_0 = x$ but $z_0 = 0$. If we now substitute $y_j = x_j - x_{j-1}$ in the numerator and $y_j = z_j - z_{j-1}$ in the denominator, we end up with

$$P_n(t)\psi(x) = \frac{\int\cdots\int \psi(x+y_1+\cdots+y_n)\exp\left(-\frac{n}{2t}\sum_{j=1}^{n} y_j^2 - \frac{t}{n}\sum_{j=1}^{n} V(x+y_1+\cdots+y_j)dy_1\cdots dy_n\right)}{\int\cdots\int \exp\left(-\frac{n}{2t}\sum_{j=1}^{n} y_j^2\right) dy_1\cdots dy_n}.$$

Now this ratio is actually an expectation of random variables – specifically, we have for all $t > 0$ that

$$P_n(t)\psi(x) = \mathbb{E}\left[\psi(x+S_n)\exp\left(-\frac{t}{n}\sum_{j=1}^{n} V(x+S_j)\right)\right],$$

where $S_j = \sum_{i=1}^{j} Y_i$ are the partial sums of iid normal random variables $Y_1, \cdots, Y_n$ of mean 0 and variance $\frac{t}{n}$. (And we define $P_n(0)$ to be the identity.) So this is something we can indeed simulate on a computer – we just need to take a large n and generate the value of the function ψ a bunch of times, taking the average of those values.

If we didn't have the V exponential term, then the expectation $\mathbb{E}[\psi(x+S_n)]$ is easy to compute because it doesn't even depend on n (since S_n is always a normal of variance t). But with that extra term, we'll need to use the fact that the S_is trace out a sample path of **Brownian motion**, and the properties of Brownian motion are what will help us do things rigorously.

Remark 9. *As $n \to \infty$, the "Lebesgue measure $dy_1\cdots dy_n$" makes less and less sense, but together with the $\exp\left(-\frac{n}{2t}\sum_{j=1}^{n} y_j^2\right)$ it will actually have a rigorous meaning. And that's what we're trying to do with all of these Yang-Mills theories too.*

Here's the exact result we're going to prove:

Theorem 10

Let $V : \mathbb{R} \to \mathbb{R}$ be continuous, bounded below, and have at most exponential growth at infinity. Then as defined above, each $P_n(t) : L^2(\mathbb{R}) \to L^2(\mathbb{R})$ is a well-defined linear operator. For each $t \geq 0$ and $\psi \in L^2(\mathbb{R})$, $P_n(t)\psi$ converges in L^2 to some $P(t)\psi$. Then the limiting operators $(P(t))_{t\geq 0}$ form a strongly continuous semigroup of linear operators on $L^2(\mathbb{R})$ with the required generator H that we desire.

(The conditions on V can be relaxed, but we won't make an effort to do so here.)

Remark 11. *Later on, we will also need "higher-dimensional analogs of Brownian motion," in the sense that the time dimension is now multidimensional (not just the spatial dimension). This will require us to construct the Gaussian free*

field, which is not even a pointwise-defined function anymore, but that construction is more complicated and thus we want to do the simpler case first.

Before we dive into this proof, we'll need to mention a few probability basics. Recall that a **probability space** is a triple $(\Omega, \mathcal{F}, \mathbb{P})$, where Ω is some set, $\mathcal{F}$ is a σ-algebra of subsets of Ω, and $\mathbb{P}$ is a probability measure on $\mathcal{F}$. Also recall that a **random variable** is a measurable function $X : \Omega \to \mathbb{R}$ (so that the inverse image of any Borel set is in $\mathcal{F}$); given any such X we can define a probability measure on $\mathbb{R}$

$$\mu_X(A) = \mathbb{P}(\{\omega \in \Omega : X(\omega) \in A\})$$

for any Borel set $A \in \mathcal{B}(\mathbb{R})$, and we call this the **law** of X. (For example, if we roll two dice and look at their sum, we usually care about the distribution of those numbers rather than of the outcomes in the original space of "dice rolls," so it's important to be careful about the subtlety of notation.) We then define the **expectation** of a random variable by saying that for any function $f : \mathbb{R} \to \mathbb{R}$, we have

$$\mathbb{E}[f(X)] = \int_\Omega f(X(\omega))dP(\omega) = \int_{\mathbb{R}} f(x)d\mu_X(x).$$

(We'll continue this a bit more next time and soon will be able to build up to what Brownian motion actually is.)

4 January 13, 2025

Today's goal is to define Brownian motion – this is a simpler version of the construction of the Gaussian free field. The idea is that instead of a real-valued random variable, we now have a function-valued random variable on (say) the time interval $[0, t]$.

We'll define Brownian motion on $[0, t]$ as a **random element of** $C([0, t])$, the space of continuous functions $[0, t] \to \mathbb{R}$. To do this, first define the sup-norm on $C([0, t])$ by $||f|| = \sup_{x\in[0,t]} |f(x)|$; this induces a distance and thus a topology on the space. Uniform limits of continuous functions are continuous, so in fact this makes $C([0, t])$ a complete separable metric space (that is, a Polish space).

This topology then defines a σ-algebra generated by the open sets (we call this the **Borel σ-algebra** of $C([0, t])$); in practice "anything that we come across" is going to be a Borel subset. A $C([0, t])$-valued random function is then a map $X : \Omega \to C([0, t])$, where $(\Omega, \mathcal{F}, \mathbb{P})$ is some probability space (so we have an experiment, and for each $\omega \in \Omega$ we associate a random curve).

Remark 12. *This is in fact a pretty practical way of thinking about these random variables – if we're trying to generate Brownian motion, our experiment gives us some random numbers, and each set of those numbers will give us a curve in some way X that we specify.*

Definition 13

A **standard Brownian motion** on $[0, t]$ is a random element of $C([0, t])$ with the following properties:

1. $B(0) = 0$,
2. (Independent increments property) For all n and all times $0 = t_0 < t_1 < \cdots < t_n \leq t$, the random variables $B(t_1 - t_0), B(t_2 - t_1), \cdots, B(t_n) - B(t_{n-1})$ are independent.
3. For all $0 \leq r < s \leq t$, the distribution of $B(s) - B(r)$ is distributed as $N(0, s - r)$.

In other words, Brownian motion is made up of independent increments, and each increment is normal with variance equal to the length of that increment. Notice that this is a rather weird condition – if r and s are quite close, then $N(0, s-r)$ is of order $\sqrt{s-r}$, which is much bigger than $s-r$. So that's what makes Brownian motion non-differentiable.

Remark 14. *Let's clarify what condition (3) above means in our more formal language. B is a map $\Omega \to C([0,t])$, so for all ω, $B(\omega)$ is a function and thus we can talk about $B(\omega)(s)$ (which is what we actually mean by $B(s)$). In other words, we can write out that condition as*

$$\mathbb{P}\left(\{\omega \in \Omega : B(\omega)(s) - B(\omega)(r) \in [a,b]\}\right) = \int_a^b \frac{1}{\sqrt{2\pi(s-r)}} \exp\left(-\frac{x^2}{2(s-r)}\right) dx.$$

Meanwhile, if we say something like $B(0) = 0$, we can take it to be $B(\omega)(0) = 0$ for all ω, but it would also be fine to only have it on a set of probability one.

The question is then whether we can explicitly construct a map $(\Omega, \mathcal{F}, \mathbb{P}) \to C([0,t])$ which satisfies these properties (so that such an object actually exists). Here's an intuitive construction with minimal assumptions: suppose we know already how to construct a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ on which we can define an infinite sequence $X_1, X_2, \cdots$ of standard normal random variables on Ω. (This is the construction of **infinite product spaces** – we can't really do this on general measurable spaces like $\mathbb{R}^\infty$, but on probability spaces it's okay because we can indeed assign measure to "cylinder sets" like $A \times B \times \Omega \times \Omega \times \cdots$). If we wanted to approximately generate Brownian motion on a computer, we'd generate a bunch of small increments, but we want to do so in a way that converges properly to a continuous curve when we refine the partition. This is **Lévy's construction**:

- First start by generating only the values $B(0)$ and $B(t)$; we know that $B(0) = 0$ and $B(t)$ should be normally distributed as $\xi(t) = N(0,t)$.
- Next, generate $B(\frac{t}{2})$; we want something chosen so that we satisfy the independent increments property. Specifically, set
$$B\left(\frac{t}{2}\right) = \frac{B(0) + B(t)}{2} + \xi\left(\frac{t}{2}\right),$$
where the new additional increment is distributed independently from $\xi(t)$ as $\xi\left(\frac{t}{2}\right) \sim N\left(0, \frac{t}{4}\right)$. Indeed $B\left(\frac{t}{2}\right)$ and $B(t)$ are jointly normal because they're linear combinations of normal random variables, and $B\left(\frac{t}{2}\right)$ is centered and has variance $\left(\frac{1}{2}\right)^2 \cdot t + \frac{t}{4} = \frac{t}{2}$. Furthermore, the increments are indeed independent because (for Gaussians we just need to check that the covariance is zero)
$$\begin{aligned}
\operatorname{Cov}\left(B(t) - B\left(\frac{t}{2}\right), B\left(\frac{t}{2}\right) - 0\right) &= \operatorname{Cov}\left(\frac{1}{2}B(t) - \xi\left(\frac{t}{2}\right), \frac{1}{2}B(t) + \xi\left(\frac{t}{2}\right)\right) \\
&= \frac{1}{4}\operatorname{Var}(B(t)) - \operatorname{Var}\left(\xi\left(\frac{t}{2}\right)\right) \\
&= \frac{t}{4} - \frac{t}{4} \\
&= 0.
\end{aligned}$$
- From here, we basically repeat the process before, iteratively defining the value of B at all dyadic rationals. (We perturb each of $B(\frac{t}{4})$ and $B(\frac{3t}{4})$ by $N(0, \frac{t}{8})$, then all odd multiples of $\frac{t}{8}$, and so on.)

More rigorously, for all j odd up to $2^k - 1$ and all $k \geq 0$, we define $X\left(\frac{j}{2^k}t\right)$ iid standard normal random variables.

(Also let $X(t)$ be another iid standard normal since it's not covered by the cases above.) Then set

$$\xi\left(\frac{j}{2^k}t\right) = \sqrt{\frac{t}{2^{k+1}}}X\left(\frac{j}{2^k}t\right).$$

We can then define $B_0(0) = 0$ and $B_0(t) = \xi(t) = X(t)$ as above, always taking linear interpolations, and then inductively define

$$B_k\left(\frac{j}{2^k}t\right) = \frac{1}{2}\left(B_{k-1}\left(\frac{j-1}{2^k}t\right) + B_{k-1}\left(\frac{j+1}{2^k}t\right)\right) + \xi\left(\frac{j}{2^k}t\right)$$

and linearly interpolating between all multiples of $\frac{1}{2^k}$.

Proposition 15

Under the sup norm, $\{B_k\}_{k\geq 0}$ is Cauchy in $C([0,t])$ with probability 1 (even though the limit itself is random and depends on the realizations of the random variables X). Thus we may define Brownian motion to be the limit of those B_ks.

Proof. Notice that because at each step we are just adding perturbations given by the ξs, we have

$$\begin{aligned}||B_k - B_{k-1}|| &= \max_{j=1,3,\cdots,2^k-1}\left|\xi\left(\frac{j}{2^k}t\right)\right| \\ &= \sqrt{\frac{t}{2^{k+1}}}\max_{j=1,3,\cdots,2^k-1}\left|X\left(\frac{j}{2^k}t\right)\right|,\end{aligned}$$

where the Xs are iid normal random variables. The idea now is that we can prove these successive distances have a finite sum, which implies that we have a Cauchy sequence, because the maximum of n standard normals is like $\sqrt{\log n}$ and thus these distances are exponentially decaying. To do this rigorously, we can use the very crude bound that for a normal random variable

$$\begin{aligned}\mathbb{P}(|X| > x) = \mathbb{P}(X^2 > x^2) &= \mathbb{P}\left(e^{X^2/4} > e^{x^2/4}\right) \\ &\leq \frac{\mathbb{E}\left[e^{X^2/4}\right]}{e^{x^2/4}} \\ &\leq Ce^{-x^2/4}\end{aligned}$$

for some constant C (since $e^{x^2/4}$ is outweighed by the Gaussian density factor, so the expectation is finite). So for n such standard normals (in fact they don't even need to be independent for this next step), we can do a union bound and get that

$$\mathbb{P}\left(\max_{1\leq i\leq n}|X_i| > x\right) \leq Cne^{-x^2/4},$$

so therefore

$$\mathbb{P}\left(||B_k - B_{k-1}|| > k\sqrt{\frac{t}{2^{k+1}}}\right) \leq \mathbb{P}\left(\max\left|X\left(\frac{j}{2^k}t\right)\right| > k\right) \leq C2^k\exp\left(-\frac{k^2}{4}\right).$$

Thus the sum of $\mathbb{P}\left(||B_k - B_{k-1}|| > k\sqrt{\frac{t}{2^{k+1}}}\right)$ over all k is finite, and thus we can apply the first Borel-Cantelli lemma (which says that if $A_1, A_2, \cdots$ is a sequence of events, aka measurable subsets, with $\mathbb{P}(A_i) < \infty$, then A_i only happens finitely often with probability one. Therefore with probability one only finitely many of the $||B_k - B_{k-1}||$s are big, and the remaining ones are at most the corresponding $k\sqrt{\frac{t}{2^{k+1}}}$ (which has a finite sum); this means the whole sum is finite and thus we do have a Cauchy sequence with probability one. This is what we are trying to prove. □

We can indeed try sampling this on a computer ourselves; we'll see that indeed we get convergence and thus can

properly sample Brownian motion! In particular, notice that at each dyadic rational t, $B_k(t)$ becomes fixed after a certain value of k, and so indeed the whole curve will stabilize. Thus the limiting curve will also have the independent increments property at dyadic rationals, and thus we can prove (by taking limits) that independent increments also hold for all real numbers in $[0, t]$.

Remark 16. *Just having exponential tails would have been good enough for our bounds, but as long as we have finite variance we'll always end up with Brownian motion (just like in the central limit theorem). So for something different we need something heavy-tailed, and then we get various processes called "stable processes" which may even be discontinuous.*

5 January 15, 2025

We'll return now to the setting from last week, in which we have a potential $V : \mathbb{R} \to \mathbb{R}$ which is continuous, bounded below, and has at most exponential growth at infinity. Recall that we're putting this problem in the context of the operators

$$P_n(t)\psi(x) = \mathbb{E}\left[\psi(x + S_n)\exp\left(-\frac{t}{n}\sum_{j=1}^{n} V(x + S_j)\right)\right]$$

for $S_j = \sum_{i=1}^{j} X_i$ the partial sums of iid normal random variables of variance $\frac{t}{n}$. (This is our approximation to the semigroup $P(t)$ that we're trying to construct, and our goal is to check that this is a semigroup of symmetric operators whose generator will then be H.)

First, we need to check that $P_n(t)$ actually maps $L^2(\mathbb{R})$ to $L^2(\mathbb{R})$. But this is simple – we just have to use the fact that $|\mathbb{E}[Z]|^2 \leq \mathbb{E}[|Z|^2]$ for any complex-valued random variable Z. We then find that

$$\begin{aligned} ||P_n(t)\psi||_{L^2}^2 &= \int_{-\infty}^{\infty} |P_n(t)\psi(x)|^2 dx \\ &\leq \int_{-\infty}^{\infty} \mathbb{E}\left[\left|\psi(x + S_n)\exp\left(-\frac{t}{n}\sum_{j=1}^{n} V(x + S_j)\right)\right|^2\right] dx, \end{aligned}$$

and now because V is bounded below by some constant we can bound this expression by $C\int_{-\infty}^{\infty} \mathbb{E}\left[|\psi(x + S_n)|\right]^2 dx$ for some constant C. By Tonelli's theorem we can now exchange the integral (because the integrand is nonnegative) and find that this is equal to

$$C\mathbb{E}\left[\int_{-\infty}^{\infty} |\psi(x + S_n)|^2 dx\right],$$

but the integral inside is just $||\psi||_{L^2}^2$ by a change of variable and thus this quantity is finite for any $\psi \in L^2$, as desired.

From here, our next step is the following:

Lemma 17

For any $\psi \in L^2(\mathbb{R})$, we have $P_n(t)\psi \to P(t)\psi$ in L^2, and the operator P is defined by

$$P(t)\psi = \mathbb{E}\left[\psi(x + B(t))\exp\left(-\int_0^t V(x + B(s))ds\right)\right]$$

where $B(t)$ is a standard Brownian motion.

We can show that $P(t)$ is also a well-defined map from L^2 to L^2, and so this justifies the Euclidean path-integral approach (justifying the expression of $P(t)$ as the thing we're aiming for).

Proof. The idea is that we can extract our iid normal random variables from the Brownian motion. If we start from a fixed Brownian motion B, then for any n we can define the random variables $X_{n,1}, \cdots, X_{n,n}$ via

$$X_{n,j} = B\left(\frac{jt}{n}\right) - B\left(\frac{(j-1)t}{n}\right),$$

and now by the defining properties the $X_{n,i}$s will be iid and distributed as $N(0, \frac{t}{n})$. So then defining $S_{n,j} = \sum_{i=1}^{n} X_{n,i}$, we thus have

$$P_n(t)\psi(x) = \mathbb{E}\left[\psi(x + S_{n,n}) \exp\left(-\frac{t}{n}\sum_{j=1}^{n} V(x + S_{n,j})\right)\right].$$

We can then compute (remember that $S_{n,n}$ is the sum of all increments of the Brownian motion and thus is **exactly** equal to $B(t)$, and similarly $S_{n,j} = B(\frac{jt}{n})$)

$$\begin{aligned}
||P_n(t)\psi - P(t)\psi||_{L^2}^2 &= \int_{-\infty}^{\infty} |P_n(t)\psi(x) - P(t)\psi(x)|^2 dx \\
&= \int_{-\infty}^{\infty} \left|\mathbb{E}\left[\psi(x + B(t))\left(\exp\left(-\frac{t}{n}\sum_{j=1}^{n} V(x + S_{n,j})\right) - \exp\left(-\int_0^t V(x + B(s))ds\right)\right)\right]\right|^2 dx \\
&\leq \int_{-\infty}^{\infty} \mathbb{E}\left[\left|\psi(x + B(t))\left(\exp\left(-\frac{t}{n}\sum_{j=1}^{n} V(x + S_{n,j})\right) - \exp\left(-\int_0^t V(x + B(s))ds\right)\right)\right|^2\right] dx.
\end{aligned}$$

But now we can use the **dominated convergence theorem** – in other words, it suffices to show that (1) for any x, the integrand goes to zero with probability 1, and (2) this integrand

$$\left|\psi(x + B(t))\left(\exp\left(-\frac{t}{n}\sum_{j=1}^{n} V(x + S_{n,j})\right) - \exp\left(-\int_0^t V(x + B(s))ds\right)\right)\right|^2$$

is bounded by some function $f(x, B)$ of x and the Brownian motion (which doesn't depend on n) with $\int_{-\infty}^{\infty} \mathbb{E}[f(x, B)]dx < \infty$. For part (1), the only part of the integrand that depends on n is the difference of exponentials. That difference $\exp\left(-\frac{t}{n}\sum_{j=1}^{n} V(x + S_{n,j})\right) - \exp\left(-\int_0^t V(x + B(s))ds\right)$ approaches zero as $n \to \infty$ because V and B are both continuous (because with probability one, a realization of the Brownian motion is continuous, so the composition of V and B is continuous, and thus the Riemann sum converges to the integral). And part (2) is trivial, because if $V(x) \geq c$ for some constant $c \in \mathbb{R}$, this difference of exponentials is at most $2\exp(-ct)$ by the triangle inequality, so we can take our function to be $|\psi(x + B(t)) \cdot 2e^{-ct}|^2$. Indeed, integrating this over B and x yields something finite, since integrating first over x (again using Tonelli's theorem) shows that we just get a constant times the L^2 norm of ψ. □

So we have now done a fully rigorous construction of the Euclidean path integral using probability – this is called the **Feynman-Kac formula**, and all we needed is that we can construct $B(t)$ and that it is continuous. But we still need to show that $P(t)$ satisfies all of our conditions (that it is a strongly continuous semigroup of symmetric operators and that it gives us the H we want).

Lemma 18

For each $t > 0$, $P(t) : L^2(\mathbb{R}) \to L^2(\mathbb{R})$ is a symmetric operator (meaning that $\langle P(t)\psi, \phi\rangle = \langle \psi, P(t)\phi\rangle$ for all $\psi, \phi \in L^2$).

Proof. Recall that ψ, ϕ are complex-valued functions, and thus the inner product is written as

$$\begin{aligned}\langle P(t)\psi, \phi\rangle &= \int_{-\infty}^{\infty} P(t)\psi(x)\overline{\phi(x)}dx \\ &= \int_{-\infty}^{\infty} \mathbb{E}\left[\psi(x+B(t))\exp\left(-\int_0^t V(x+B(s))ds\right)\overline{\phi(x)}\right]dx.\end{aligned}$$

We can exchange the expectation and integral by Fubini's theorem (the integrand is integrable), so we're left with

$$\mathbb{E}\left[\int_{-\infty}^{\infty}\psi(x+B(t))\overline{\exp\left(-\int_0^t V(x+B(s))ds\right)\phi(x)}dx\right]$$

where we use that V is real-valued, so it's okay to put a bar over the exponential term. Making a change of variables $y = x + B(t)$ turns this into

$$\mathbb{E}\left[\int_{-\infty}^{\infty}\psi(y)\overline{\exp\left(-\int_0^t V(y+B(s)-B(t))ds\right)\phi(y-B(t))}dy\right].$$

If we now define a new stochastic process $W : [0,t] \to \mathbb{R}$ by $W(s) = B(t-s) - B(t)$, then $W(0) = 0$ and $W(t) = -B(t)$ – this is essentially the **time-reversal** of B. We can check that this is a continuous Gaussian process (any finitely many points are jointly normal), $\mathbb{E}[W(s)] = 0$ for all $s \in [0,t]$, and it has the same covariance structure as Brownian motion (meaning that it also has independent increments and $W(s) - W(r) \sim N(0, s-r)$); in particular, it must still be a Brownian motion because the defining properties give us a unique law. So our expression actually becomes

$$\mathbb{E}\left[\int_{-\infty}^{\infty}\psi(y)\overline{\exp\left(-\int_0^t V(y+W(s))ds\right)\phi(y+W(t))}dy\right].$$

So applying Fubini's theorem again and moving the expectation back inside yields

$$\int_{-\infty}^{\infty}\psi(y)\overline{\mathbb{E}\left[\phi(y+W(t))\exp\left(-\int_0^t V(y+W(s))ds\right)dy\right]} = \int_{-\infty}^{\infty}\psi(y)\overline{P(t)\phi(y)}dy,$$

and this is exactly the other inner product $\langle\psi, P(t)\phi\rangle$ that we were looking for, as desired. □

Lemma 19

For all $s, t > 0$, we have $P(s+t) = P(s)P(t)$.

Proof. Let W be Brownian motion, so that we have

$$\begin{aligned}P(s)\phi(x) &= \mathbb{E}\left[\phi(x+W(s))\exp\left(-\int_0^s V(x+W(u))du\right)\right] \\ &= \int_{C([0,s])}\phi(x+w(s))\exp\left(-\int_0^s V(x+w(u))du\right)d\mu(w),.\end{aligned}$$

where w is the law of Brownian motion on $C([0,s])$. (This is what we were talking about a few lectures ago where expectations can be evaluated relative to the original probability space as well.) If we now define $\phi = P(t)\psi$, we have (by the same reasoning) that

$$\phi(y) = \int_{C([0,t])}\psi(y+b(t))\exp\left(-\int_0^t V(y+b(u))du\right)d\nu(b),$$

where ν is now the law of Brownian motion on $C([0,t])$. So substituting one expression into the other, we find that

$$P(s)P(t)\psi(x) = \int_{C([0,s])} \left(\int_{C([0,t])} \psi(x + w(s) + b(t)) \exp\left(- \int_0^t V(x + w(s) + b(u))du \right) d\mu(w) \right) \exp\left(- \int_0^s V(x + w(u))du \right) d\nu(b),$$

which we can then rewrite more cleanly as an expectation involving two Brownian motions

$$\int_{C([0,s])} \int_{C([0,t])} \psi(x + w(s) + b(t)) \exp\left(- \int_0^t V(x + w(s) + b(u))du \right) \exp\left(- \int_0^s V(x + w(u))du \right) d\mu(w) d\nu(b).$$

Putting this back in the original expectation format, this is thus

$$\mathbb{E}\left[\psi(x + W(s) + B(t)) \exp\left(- \int_0^t V(x + W(s) + B(u))du - \int_0^s V(x + W(u))du \right) \right]$$

where W and B are independent Brownian motions on $[0,s]$ and $[0,t]$. So we can now define the process which puts together those two Brownian motions into a single one:

$$Z(u) = \begin{cases} W(u) & 0 \le u \le s, \\ W(s) + B(u-s) & s \le u \le t \end{cases},$$

and then the expression above exactly becomes $\mathbb{E}\left[\psi(x + Z(s+t)) \exp\left(- \int_0^{s+t} V(x + Z(u))du \right) \right]$, which is exactly $P(t+s)\psi(x)$ as desired. □

In summary, we can think of this as being a "Markov property:" if we look at Brownian motion beyond time s, the distribution given the past only depends on the current endpoint $B(s)$. And now it just remains to show strong continuity (easy) and that we have the correct generator.

Remark 20. *We'll do all of these things with the Euclidean free field soon; the only treatment of this material that is needed for our constructions is found in some old books, but those are quite old and don't quite say things in the probability language.*

6 January 17, 2025

Last time, we proved that our collection $(P(t))_{t \ge 0}$ of linear operators on $\mathcal{H} = L^2(\mathbb{R})$ is indeed a semigroup. To prove strong continuity now, we'll use the following useful lemma:

Lemma 21

Let $\mathcal{H}$ be a Hilbert space, and let $(P(t))_{t \ge 0}$ be any collection of linear oprators from $\mathcal{H}$ into $\mathcal{H}$. Suppose that

1. for all $T \ge 0$, $\sup_{0 \le t \le T} ||P(t)|| < \infty$, and
2. there is some subset $S \subseteq \mathcal{H}$ with dense span (that is, $\overline{\text{span}(S)} = \mathcal{H}$, where the span is the set of all finite linear combinations of elements in S), and $t \mapsto P(t)x$ is continuous for all $x \in S$.

Then $t \mapsto P(t)x$ is in fact continuous for all $x \in \mathcal{H}$.

In other words, we only need to verify strong continuity on a smaller subset, as long as we have some control over the norm of P.

Proof. Consider any $x \in \mathcal{H}$ and $t \geq 0$, and let $t_n \to t$ be a sequence of nonnegative real numbers. We must show that $P(t_n)x \to P(t)x$. First of all, there is some T with $t_n \in [0, T]$ for all n (because the sequence converges to t); let $C = \sup_{0 \leq t \leq T} ||P(t)||$. By assumption, for any $\varepsilon > 0$, there is some $y \in \text{span}(S)$ such that $||x - y|| < \varepsilon$. We then do the usual three-term bound

$$\begin{aligned} ||P(t_n)x - P(t)x|| &\leq ||P(t_n)x - P(t_n)y|| + ||P(t_n)y - P(t)y|| + ||P(t)y - P(t)x|| \\ &\leq C||x - y|| + ||P(t_n)y - P(t)y|| + C||x - y|| \\ &< 2C\varepsilon + ||P(t_n)y - P(t)y||, \end{aligned}$$

and as $n \to \infty$ the latter term goes to zero (because y is a finite linear combination of elements s for which $P(t_n)s \to P(t)s$) so $\limsup ||P(t_n)x - P(t)x|| \leq 2C\varepsilon$ and therefore we must actually have $||P(t_n)x - P(t)x|| \to 0$. $\square$

We'll use this theorem later on several times when dealing with quantum fields as well, but let's apply it to our operators here first. We have

$$P(t)\psi(x) = \mathbb{E}\left[\psi(x + B(t)) \exp\left(-\int_0^t V(x + B(s))ds\right)\right]$$

and thus by the same kinds of bounds as before (and swapping the order of integrals), we have

$$\begin{aligned} ||P(t)\psi||_{L^2}^2 &\leq \int_{-\infty}^{\infty} \mathbb{E}\left[\left|\psi(x + B(t)) \exp\left(-\int_0^t V(x + B(s))ds\right)\right|\right]^2 \\ &\leq e^{-2ct} \int_{-\infty}^{\infty} \mathbb{E}\left[|\psi(x + B(t))|\right]^2 \\ &\leq e^{-2ct}||\psi||_{L^2}^2, \end{aligned}$$

where c is the constant that V is bounded by from below. This may exponentially explode as $t \to \infty$ because c may be negative, but this is indeed uniformly bounded on $[0, T]$ for any T. (And this is pretty close to the best bound we can get, since remember that $P(t) = e^{-tH}$ and then the norm depends on the smallest eigenvalue of H, which may be negative.)

To verify the second condition, we'll now check that $t \mapsto P(t)\psi$ is continuous **when ψ is continuous with compact support** (which is indeed a dense subset of $L^2(\mathbb{R})$). We have for any s, t (by a similar calculation as before again) that

$$||P(s)\psi - P(t)\psi||_{L^2}^2 \leq \int_{-\infty}^{\infty} \mathbb{E}\left[\left|\psi(x + B(s)) \exp\left(-\int_0^s V(x + B(u))du\right) - \psi(x + B(t)) \exp\left(-\int_0^t V(x + B(u))du\right)\right|^2\right].$$

As $s \to t$, the whole expression inside the absolute values approaches 0 as $s \to t$ because ψ, B, and V are all continuous functions. So we just need to verify the other condition for applying the dominated convergence theorem – we need to show this is dominated by a random variable that doesn't depend on s whose expectation is finite. Notice that we can't just bound ψ by a constant overall, because the bounded convergence theorem works on probability spaces but here we're integrating x on all of $\mathbb{R}$. But because ψ has compact support, say contained in $[-L, L]$ for some L, we know that our integrand is zero if $|x + B(s)| > L$ and $|x + B(t)| > L$. In other words, we get a bound (here C

accounts for both the magnitude of ψ and of the bound coming from V)

$$\begin{aligned}\left|\psi(x+B(s))\exp\left(-\int_0^s V(x+B(u))du\right)-\psi(x+B(t))\exp\left(-\int_0^t V(x+B(u))du\right)\right|^2 \\ \leq C\cdot 1\{|x+B(s)|\leq L, |x+B(t)|\leq L\} \\ \leq C\cdot 1\{|x|\leq |B(s)|+L, |x|\leq |B(t)|+L\} \\ \leq C\cdot 1\{|x|\leq M+L\},\end{aligned}$$

where M is the maximum of $|B(u)|$ on the interval $[0, t+1]$ (we consider s approaching t, so we can assume that it's close enough to t that it's in this interval). But now the expectation of this last indicator is indeed finite, since

$$\int_{-\infty}^{\infty}\mathbb{E}\left[1\{|x|\leq M+L\}\right]dx=\mathbb{E}\int_{-\infty}^{\infty}\left[1\{|x|\leq M+L\}\right]dx=\mathbb{E}[2(M+L)],$$

and the maximum of Brownian motion on a finite interval has finite expectation (by some direct calculations). Thus the dominated convergence theorem applies and as $s\to t$ this will converge to zero, as desired.

Recall that we already have the existence of a self-adjoint operator H in our example, and what we're doing here is checking that the corresponding generator for $P(t)$ is also that same H:

Lemma 22

If ψ is C^∞ (smooth) with compact support, then we have the L^2 limit

$$\lim_{t\downarrow 0}\frac{P(t)\psi-\psi}{-t}=H\psi,$$

where $H\psi=-\frac{1}{2}\psi''+V\psi$.

Proof. Let $\psi_t=P(t)\psi$, so that $\psi_0=\psi$. We'll use an approximate process: first we approximate the exponential via

$$\xi_t(x)=(1-tV(x))\mathbb{E}\left[\psi(x+B(t))\right],$$

and then we further approximate with a Taylor expansion

$$\eta_t(x)=(1-tV(x))\left(\psi(x)+\frac{1}{2}\psi''(x)t\right)$$

We will show that $\frac{\psi_t-\xi_t}{-t}\to 0$ and $\frac{\xi_t-\eta_t}{-t}\to 0$ both in L^2, which will complete the proof (since if we take the corresponding limit $\frac{\eta_t\psi-\psi}{-t}$ as $t\downarrow 0$, we indeed get exactly H).

This is a fairly routine check, so we won't do all the details: we have

$$\psi_t(x)-\xi_t(x)=\mathbb{E}\left[\psi(x+B(t))\left(\exp\left(-\int_0^t V(x+B(s))ds\right)-(1-tV(x))\right)\right],$$

which means that (with the same familiar tools)

$$\int_{-\infty}^{\infty}|\psi_t(x)-\xi_t(x)|^2dx\leq\int_{-\infty}^{\infty}\mathbb{E}\left[|\psi(x+B(t))|^2\exp\left(-\int_0^t V(x+B(s))ds-(1-tV(x))\right)^2\right]dx,$$

and moving the expectation outside the integral and doing the change of variables $y=x+B(t)$ yields

$$\mathbb{E}\left[\int_{-\infty}^{\infty}|\psi(y)|^2\left(\exp\left(-\int_0^t V(y+B(s)-B(t))ds\right)-1+tV(y-B(t))\right)^2\right]$$

The integrand goes to zero as $t \to 0$ by Taylor expansion, and since ψ is zero outside some finite integral, then even after dividing by t we see that this whole expectation goes to zero (again this involves some maximum of Brownian motion on a compact interval).

Finally, for the other term, we wish to consider

$$(1 - tV(x))\mathbb{E}\left[\psi(x + B(t))\right] - (1 - tV(x))(\psi(x) + \frac{t}{2}\psi''(x)).$$

We may wonder what happened to the first derivative of ψ; the point is that

$$\mathbb{E}[\psi(x + B(t))] = \mathbb{E}\left[\psi(x) + B(t)\psi'(x) + \frac{1}{2}B(t)^2\psi''(x) + \cdots\right],$$

but the expectation of $B(t)$ is zero and the expectation of $B(t)^2$ is t, and everything else is higher-order. And then the whole thing can be made rigorous by the dominated convergence theorem and more reasoning about Brownian motion. □

Ultimately, what all of this gives us is that we have a self-adjoint extension of the Schrodinger operator for a one-dimensional potential V of the restrictions we've assigned, and so by Stone's theorem we can construct solutions to the Schrodinger equation. (This is quite nontrivial directly even for something like $V(x) = x^4$.) But the point is that this whole program also gives us a strategy with quantum fields as well.

Remark 23. *Given a Euclidean field theory, there's a set of conditions called the "Osterwalder–Schrader axioms" that let us get the quantum field theory. But that's the same amount of work, and for any new system we always have to go through this whole process to do a rigorous construction. And our example here has been quite simple, because we've just been dealing with $L^2(\mathbb{R})$ rather than a more complicated Hilbert space.*

We'll begin with the simplest examples of fields:

Definition 24

A **classical Klein-Gordon field** is a function $\phi : \mathbb{R}^4 \to \mathbb{R}$ evolving via the equation (here our four coordinates are t, x, y, z)

$$\frac{\partial^2\phi}{\partial t^2} = \Delta\phi - m^2\phi,$$

where $m \geq 0$ (in particular, $m = 0$ will later correspond to the free electromagnetic field) and where $\Delta\phi = \frac{\partial^2\phi}{\partial x^2} + \frac{\partial^2\phi}{\partial y^2} + \frac{\partial^2\phi}{\partial z^2}$.

These solutions ϕ are critical points of a certain action

$$S(\phi) = \int_{\mathbb{R}^4} \left((\partial_t\phi)^2 - |\nabla\phi|^2 - m^2\phi^2\right) dtdxdydz.$$

In other words, if for any perturbation of compact support for ϕ, the variational derivative of S must vanish, then ϕ solves the Klein-Gordon equation. And the point is that once we have this action, we know how to quantize via path integrals – we have a field, and thus we have a **wavefunction on the space of fields** (so this is kind of like having a one-dimensional particle at every point in space) whose square tells us the probability density on the space of fields. So what we'll do is write down the action, write down the Wick action, and study the Euclidean action, just like we did with a single particle.

7 January 22, 2025

Last time, we introduced the Klein-Gordon equation $\frac{\partial^2\phi}{\partial t^2} = \Delta\phi - m^2\phi$ for a field evolving in time – for $m = 0$ we get the wave equation, but in general m represents the mass of the particle. We also claimed that solutions to this equation are critical points of a particular action

$$S(\phi) = \int_{\mathbb{R}^4} \left((\partial_t\phi)^2 - |\nabla\phi|^2 - m^2\phi^2\right) dtdx.$$

So this is a simpler version of Maxwell's equations for electromagnetic waves; we'll use the techniques we develop here to do a quantization of the free electromagnetic field later.

The path integral quantization of the Klein-Gordon field will be the "free quantum field." To understand what that means, note that ϕ gives us a height function defined on $\mathbb{R}^3$ at each point in time, and then we have some quantum evolution of ϕ. States of our system at a given time are then rays of a Hilbert space $\mathcal{H}$ (so we must first figure out what $\mathcal{H}$ is), and ideally we would like $\mathcal{H}$ to be the set of all L^2 functions on some space $\mathcal{X}$ of "all possible values of $\phi(t,\cdot)$." A priori it's not clear what space to take – from our previous discussion, we might expect that it consists of very rough functions, and it turns out that it will actually consist of very rough distributions. Indeed, given some initial state $\psi_0 \in \mathcal{H}$ (which inputs an entire field and outputs a complex number), our state should evolve to some ψ_t at a later time t. This is characterized by (for any particular field $\phi_* \in \mathfrak{X}$ as input)

$$\psi_t(\phi_*) = \frac{1}{Z(t)} \int_{\phi:\phi(0,\cdot)=\phi_*} \psi_0(\phi(t,\cdots))e^{iS(\phi)}\mathcal{D}\phi,$$

where the action is $S(\phi) = \int_0^t \int_{\mathbb{R}^3}((\partial_s\phi)^2 - |\nabla\phi|^2 - m^2\phi^2)dsdx$. But we need to figure out what $\mathfrak{X}$ is, how to get a Lebesgue measure on it, and then what the definition of the evolution ψ_t actually is.

Example 25

As before, we'll first do a Wick rotation to turn this into a Euclidean evolution, and we'll try to show that the semigroup has the necessary properties to get a Hamiltonian.

(Notice that we don't know what the Hamiltonian is this time – in the Schrodinger example, we verified that the Hamiltonian we got was actually the right one, but here we'll do the Wick rotation, get a $P(t)$ and H, and we'll take that H **as the definition**.)

If we plug in $t \mapsto -it$ in the above evolution, we get the new evolution

$$P(t)\psi(\phi_*) = \frac{1}{\tilde{Z}(t)} \int_{\phi:\phi(0,\cdot)=\phi_*} \psi(\phi(t,\cdot))e^{-S_E(\phi)}\mathcal{D}\phi,$$

with (details with how the factors of i show up in the rest of the expression omitted, but we can check that it all works out)

$$S_E(\phi) = \int_0^t \int_{\mathbb{R}^3} \left((\partial_s\phi)^2 + |\nabla\phi|^2 + m^2\phi^2\right) dsdx.$$

Notice that we previously pick Z so that we get a unitary evolution, and now we pick $\tilde{Z}$ so that $P(t)$ is a symmetric strongly continuous semigroup. (The problem with getting an expression for $\tilde{Z}$ is that the mass term makes this not invariant under shifting the field by a constant, but the expression above depends on the starting state ϕ_*; we don't want the normalizing constant to depend on the argument that we put in. We'll come back to this later in the course.) But the point is that this motivates the following construction (for which we care about the $d = 4$ case):

Definition 26

Let $d \geq 3$ and $m \geq 0$. The **Euclidean free field on $\mathbb{R}^d$ with mass m** is formally a real-valued Gaussian field $\phi : \mathbb{R}^d \to \mathbb{R}$ with probability density proportional to

$$\exp\left(-\frac{1}{2}\int_{\mathbb{R}^d} \left(|\nabla\phi|^2 + m^2\phi^2\right) dx\right)$$

with respect to "the Lebesgue measure on the space of all ϕ."

This expression comes from the expression in the integrand of the Euclidean action S_E – time and space are completely symmetric in that action, so we can just do everything symmetrically over d coordinates. And from here, the evolution will come from looking at time-slices of our distribution; we get a Markov process and $P(t)$ will be the semigroup of that Markov process. (Notice that we cannot actually construct this object when $d = 1, 2$ and $m = 0$; we'll see why in a minute.)

If we define for real-valued functions ϕ, ψ (we won't worry about the domain for now)

$$Q(\phi, \psi) = \int \left(\nabla\phi \cdot \nabla\psi + m^2\phi\psi\right) dx,$$

then this is formally a positive definite inner product, and our goal is to get a probability density on the function space which looks like $\exp\left(-\frac{1}{2}Q(f, f)\right)$, which is exactly what we get with Gaussians. So this indicates that we should have a centered Gaussian field whose covariance structure should be given by Q^{-1}, since by integration by parts

$$Q(f, g) = \int_{\mathbb{R}^d} f \cdot (-\Delta + m^2 I)g \, dx$$

and so the covariance of ϕ should be given by the inverse $Q^{-1} = (-\Delta + m^2 I)^{-1}$. If instead we had some other kernel K with $Kg(x) = \int K(x, y)g(y)dy$, then we'd be exactly in the familiar Gaussian field density situation, but this operator $-\Delta + m^2 I$ is not an integral kernel. It turns out the inverse of this operator is though – we can define for any $x \neq 0$

$$G_m(x) = \int_0^\infty \frac{1}{(4\pi t)^{d/2}} \exp\left(-\frac{||x||^2}{4t} - m^2 t\right) dt,$$

and then for any **Schwartz function** f (that is, an infinitely differentiable function where it and all of its functions are **rapidly decaying**, meaning they decay faster than any polynomial) we define the operator

$$G_m f(x) = \int_{\mathbb{R}^d} G_m(x - y)f(y)dy,$$

so that we have $(-\Delta + m^2 I)G_m f = G_m(-\Delta + m^2 I) = f$. This means we have an inverse G_m (and here is where we need $d \geq 3$ for $m = 0$ so that the integral expression does converge), and we have a candidate object such that the feature

$$\mathrm{Cov}(\phi(x), \phi(y)) = G_m(x - y)$$

should hold. But from this, we see that the variance of $\phi(x)$ and $\phi(y)$ tends to infinity as $x \to y$; thus we'll define ϕ as a distribution instead, specifically specifying a collection $\{\phi(f) : f \in \mathcal{S}(\mathbb{R}^d)\}$ of jointly normal random variables such that $\mathrm{Cov}(\phi(f), \phi(g)) = \iint f(x)g(y)G_m(x - y)dxdy$. (Here we can interpret $\phi(f)$ as $\int_{\mathbb{R}^d} \phi(x)f(x)dx$.)

There are much more refined things we can do such as asking for continuity in f, but we won't go into that – we'll just define it as a collection of Gaussian random variables and work with the characterization (similarly to defining Brownian motion via the Kolmogorov extension theorem).

8 January 24, 2025

We've been working on constructing the Euclidean free field in dimension $d \geq 3$ by defining the function $G_m : \mathbb{R}^d \to \mathbb{R}$ via

$$G_m(x) = \int_0^\infty \frac{1}{(4\pi t)^{d/2}} \exp\left(-\frac{||x||^2}{4t} - m^2 t\right) dt$$

for any $x \neq 0$ and then setting $G_m(0) = \infty$. When we have $m = 0$, this expression is easy to evaluate: letting $s = \frac{||x||^2}{4t}$, we find that $t = \frac{||x||^2}{4s} \implies dt = -\frac{||x||^2}{4s^2} ds$, so that $t^{d/2} = \frac{||x||^d}{(4s)^{d/2}}$. Plugging everything back in, we thus find that

$$\begin{aligned} G_0(x) &= \int_0^\infty \frac{(4s)^{d/2}}{||x||^d (4\pi)^{d/2}} e^{-s} \frac{||x||^2}{4s^2} ds \\ &= \frac{1}{||x||^{d-2}} \int_0^\infty \frac{s^{d/2-2}}{4\pi^{d/2}} e^{-s} ds \\ &= \frac{\Gamma(d/2-1)}{4\pi^{d/2}||x||^{d-2}}. \end{aligned}$$

In particular, we can see from this explicit form that this blows up near $x = 0$ but decays near infinity; by the same technique, we get a lower bound $G_m(x) \leq \frac{C(d)}{||x||^{d-2}}$ for all m (since we can always just ignore the $-m^2$ term in writing down an inequality). It turns out that we actually get better behavior as $x \to \infty$ (in fact exponential decay), though we will not need that.

Lemma 27

G_m sends Schwartz functions to C^∞ functions, and it commutes with derivatives.

Proof. This is an application of the dominated convergence theorem. Recall that we define the operator G_m via

$$\begin{aligned} G_m f(x) &= \int_{\mathbb{R}^d} G_m(x-y) f(y) dy \\ &= \int_{\mathbb{R}^d} G_m(y) f(x-y) dy. \end{aligned}$$

By the decay properties of f and its derivatives (from the definition of being a Schwartz function), it is now easy to show that $\partial_i G_m f = G_m \partial_i f$. (To show that the last integral is well-defined, we can break it up into a sum over $|y| \leq 1$ and another sum where $|y| \geq 1$; in the latter term we can use our bound on G_m and in the former we can just bound f by a constant and show that the integral $\int_{||y||\leq 1} \frac{1}{||y||^{d-2}} dy$ is finite. Thus there is no problem.)

That fact is sufficient for completing the proof, since we can now do this argument repeatedly (using the fact that all derivatives of a Schwartz function are themselves Schwartz). Thus $G_m f$ is in C^∞ as desired. □

Notice that the integral $\int_{||y||\leq 1} \frac{1}{||y||^{d-1}} dy$ is also finite, and later we'll see that this will be necessary for time-slices of the free field to make sense. And in general, whether we can use a certain Green's function will depend on the class of test functions we use – it was important here that f has some decay behavior.

Theorem 28

For any Schwartz function $f \in \mathcal{S}(\mathbb{R}^d)$, we have

$$(-\Delta + m^2 I) G_m f = G_m (-\Delta + m^2 I) f = f.$$

This actually requires quite a bit of work to show (though there may be other ways to prove this using Fourier transforms)

Proof. Let $g = G_m f$. Then we know that $g \in C^\infty(\mathbb{R}^d)$, and by commutativity with derivatives we have $\Delta g = \Delta G_m f = G_m \Delta f$. This expression evaluates out to

$$\begin{aligned}\int_{\mathbb{R}^d} G_m(x-y)\Delta f(y)dy &= \int_{\mathbb{R}^d}\int_0^\infty \frac{1}{(4\pi t)^{d/2}}\exp\left(-\frac{||x-y||^2}{4t} - m^2 t\right)\Delta f(y)dtdy \\ &= \int_0^\infty \int_{\mathbb{R}^d} \frac{1}{(4\pi t)^{d/2}}\exp\left(-\frac{||x-y||^2}{4t} - m^2 t\right)\Delta f(y)dydt\end{aligned}$$

(interchanging the order of integrals by Fubini's theorem; note that this is particularly important when our functions have singularities). Now for any $t > 0$, we have

$$\left|\int_{\mathbb{R}^d} \frac{1}{(4\pi t)^{d/2}}\exp\left(-\frac{||x-y||^2}{4t} - m^2 t\right)\Delta f(y)dy\right| \leq ||\Delta f||_{L^\infty}\int_{\mathbb{R}^d}\frac{1}{(4\pi t)^{d/2}}\exp\left(-\frac{||x-y||^2}{4t}\right),$$

and that integral on the right-hand side is just 1. Thus we can use the dominated convergence theorem to write the previous quantity as

$$G_m\Delta f = \lim_{\varepsilon\to 0}\int_\varepsilon^\infty\left(\int_{\mathbb{R}^d}\frac{1}{(4\pi t)^{d/2}}\exp\left(-\frac{||x-y||^2}{4t} - m^2 t\right)\Delta f(y)dy\right)dt.$$

Looking at this inner integral more carefully now (except for the $\exp(-m^2 t)$, which we can bring outside), we have by integration by parts that

$$\begin{aligned}\int_{\mathbb{R}^d}\exp\left(-\frac{||x-y||^2}{4t}\right)\Delta f(y)dy &= \int_{\mathbb{R}^d} f(y)\Delta_y\exp\left(-\frac{||x-y||^2}{4t}\right)dy \\ &= \int_{\mathbb{R}^d} f(y)\left(-\frac{d}{2t} + \frac{||x-y||^2}{4t}\right)\exp\left(-\frac{||x-y||^2}{4t}\right)dy.\end{aligned}$$

Substituting this back in and using Fubini's theorem again, we thus get

$$G_m\Delta f(x) = \lim_{\varepsilon\to 0}\int_{\mathbb{R}^d}\int_\varepsilon^\infty f(y)\left(-\frac{d}{2t} + \frac{||x-y||^2}{4t}\right)\exp\left(-\frac{||x-y||^2}{4t}\right)\frac{e^{-m^2 t}}{(4\pi t)^{d/2}}dtdy.$$

And now we see why we use this specific Green's function: taking a time-derivative, we actually see that $\frac{\partial}{\partial t}\left(\frac{1}{(4\pi t)^{d/2}}\exp\left(-\frac{||x-y||^2}{4t}\right)\right) = \frac{1}{(4\pi t)^{d/2}}\left(-\frac{d}{2t} + \frac{||x-y||^2}{4t}\right)\exp\left(-\frac{||x-y||^2}{4t}\right)$ (This is what it means to be the fundamental solution of the heat equation.) So we can again use an integration by parts and simplify the above to

$$\begin{aligned}G_m\Delta f(x) &= \lim_{\varepsilon\to 0}\int_{\mathbb{R}^d}\int_\varepsilon^\infty f(y)e^{-m^2 t}\frac{\partial}{\partial t}\left(\frac{1}{(4\pi t)^{d/2}}\exp\left(-\frac{||x-y||^2}{4t}\right)\right)dtdy \\ &= \lim_{\varepsilon\to 0}\int_{\mathbb{R}^d}\left[\left(f(y)e^{-m^2 t}\frac{1}{(4\pi t)^{d/2}}\exp\left(-\frac{||x-y||^2}{4t}\right)\right)_\varepsilon^\infty + m^2\int_\varepsilon^\infty f(y)e^{-m^2 t}\frac{1}{(4\pi t)^{d/2}}\exp\left(-\frac{||x-y||^2}{4t}\right)dtdy\right].\end{aligned}$$

(Note that in this last step we had to be careful about the boundary when applying integration by parts because of $t = \varepsilon$.) Evaluating at the endpoints and taking $\varepsilon \to 0$ in the latter part yields

$$\lim_{\varepsilon\to 0}\int_{\mathbb{R}^d} -f(y)e^{-m^2\varepsilon}\frac{1}{(4\pi\varepsilon)^{d/2}}\exp\left(-\frac{||x-y||^2}{4\varepsilon}\right) + m^2\int_{\mathbb{R}^d}\int_0^\infty f(y)e^{-m^2 t}\frac{1}{(4\pi t)^{d/2}}\exp\left(-\frac{||x-y||^2}{4t}\right)dtdy,$$

so that the latter term is $m^2 G_m f = m^2 g$; the nontrivial part is showing that the first term evaluates to f. (So this boundary term does actually matter!) Doing a change of variable $z = \frac{x-y}{\sqrt{\varepsilon}}$, $dz = \varepsilon^{-d/2}dy$, that first integral becomes

$$-\lim_{\varepsilon\to 0}\int_{\mathbb{R}^d} f(x-\sqrt{\varepsilon}z)\frac{1}{(4\pi)^{d/2}}e^{-||z||^2/4}dz;$$

other than the f, this integrand is just the probability density of a normal random variable. But f is nice enough that

we can apply the dominated convergence theorem as $\varepsilon \to 0$, and so we're left with $-f$ in this expression as desired. Thus $\Delta g = -f + m^2 g$, which rearranges to $f = (-\Delta + m^2 I)g$, as desired.

The opposite property is now easy because G_m commutes with derivative operators, and thus G_m is indeed the inverse of $-\Delta + m^2 I$ on both sides, completing the proof. □

The same proof also works for $d = 1$ or $d = 2$ for m strictly positive so that G_m is well-defined. (But we're most curious about the $m = 0$ case because we're going to construct the free electromagnetic field later.)

Remark 29. *There's still a Green's function in $d = 2$ – it's logarithmic – but we can't really define the Gaussian free field with it. This is because if we take correlation between $\phi(x)$ and $\phi(y)$ in one or two dimensions, that's blowing up instead of decaying.*

We'll now describe the general construction of Gaussian fields; the problem is that they really only work in this specialized setting. For $f, g \in \mathcal{S}(\mathbb{R}^d)$, we define the correlation kernel (which is a bilinear form on Schwartz functions)

$$R(f,g) = \iint f(x)g(y)G_m(x-y)dxdy.$$

Lemma 30

R is a positive definite inner product on $\mathcal{S}(\mathbb{R}^d)$; this is the most basic thing we need to actually define a correlation structure. Furthermore, the topology induced by this inner product is separable (meaning there is a countable dense subset).

Start of proof. Since G_m is always nonnegative, we can check for integrability via the bound

$$\iint G_m(x-y)|f(x)g(y)|dxdy \le C \iint \frac{1}{||x-y||^{d/2}}|f(x)g(y)|dxdy.$$

We now split this up into parts where y is within 1 unit of x or not, and the corresponding bound is that

$$\begin{aligned}\iint G_m(x-y)|f(x)g(y)|dxdy &\le C \iint_{||y-x||\le 1} \frac{|f(x)g(y)|}{||x-y||^{d-2}}dxdy + C\iint_{||y-x||>1}\frac{|f(x)g(y)|}{||x-y||^{d-2}}dxdy \\ &\le C\int |f(x)|dx \cdot ||g||_{L^\infty} + C\iint |f(x)||g(y)|dxdy \\ &\le C\left(||f||_{L^1}||g||_\infty + ||f||_{L^1}||g||_{L^1}\right).\end{aligned}$$

This bound will help us prove positive definiteness (being a bilinear form is clear). We have

$$R(f,f) = \int_0^\infty \frac{e^{-m^2 t}}{(4\pi t)^{d/2}}\left(\iint f(x)f(y)\exp\left(-\frac{||x-y||^2}{4t}\right)dxdy\right)dt.$$

And next time, we'll show that this is nonnegative for any Schwartz function f and zero only if f is zero. □

From here, the idea is to use Gram-Schmidt on the countable subset to get an orthonormal basis, and we construct the free field by taking a sequence of iid normal variables to be the coefficients in this basis. But we'll see that next time!

Remark 31. *Notice that when we do this kind of thing, we always have a covariance and an inverse covariance. The inverse covariance is what shows up in the density, but the covariance is what we're working on here and what we need for the construction.*

9 January 27, 2025

Last time, we defined the bilinear form $R(f,g) = \iint f(x)g(y)G_m(x-y)dxdy$ which is supposed to be the covariance function for the Gaussian free field. We'll finish the proof of Lemma 30 today:

Proof of Lemma 30, continued. Last time, we proved an estimate $\iint G_m(x-y)|f(x)g(y)|dxdy \le C||f||_{L^1}||g||_{L^\infty} + C||f||_{L^1}||g||_{L^1}$; in particular this proves that the integral is at least well-defined as a Lebesgue integral for Schwartz functions. We then calculated that

$$R(f,f) = \int_0^\infty \frac{e^{-m^2 t}}{(4\pi t)^{d/2}} \left(\iint f(x)f(y) \exp\left(-\frac{||x-y||^2}{4t}\right) dxdy \right) dt,$$

and so what we want to do is prove that the inner parenthetical part is nonnegative and only zero if $f = 0$. For this, notice that

$$e^{-||x-y||^2/4t} = \mathbb{E}\left[e^{i(x-y)\cdot Z/\sqrt{2t}}\right],$$

where Z is a d-dimensional standard Gaussian; by Fubini's theorem we can bring the expectation outside the double integral and thus the blue part evaluates to

$$\mathbb{E}\left[f(x)f(y)e^{i(x-y)\cdot Z/\sqrt{2t}}\right] = \mathbb{E}\left[\left|\int f(x)e^{ix\cdot Z/\sqrt{2t}}dx\right|^2\right]$$

by factoring out the parts for x and y (and using that we get a negative sign in the exponential when we take a conjugate, and also that f is real-valued). So we indeed have nonnegativity; furthermore what we have here is the Fourier transform $\hat{f}(Z/\sqrt{2t})$, so that is zero if and only if it is zero with probability one. Thus if $R(f,f) = 0$, then the function $\hat{f}$ must be zero almost everywhere, and since it is a continuous function, f is zero as desired.

Finally, we prove separability. Recalling that $L^1(\mathbb{R}^d)$ is separable and that any subset of a separable metric space is separable (these are easy exercises for us to try), we see that for all n the set

$$\mathcal{S}_N = \{f : f \text{ Schwartz}, ||f||_{L^\infty} \le N\}$$

is separable in the L^1 topology. But because of the estimate (plugging in f for both arguments)

$$||f||_R^2 = R(f,f) \le C||f||_{L^1}||f||_{L^\infty} + C||f||_{L^1}||f||_{L^1},$$

we see that $\mathcal{S}_N$ is separable in the R-topology. Since this holds for all integers N, this means $\mathcal{S}(\mathbb{R}^d)$, a countable union of separable sets, is separable in the R-topology as well. This completes the proof. □

So now we can take any countable dense subset of $\mathcal{S}(\mathbb{R}^d)$ (in the R-topology), and apply Gram-Schmidt orthogonalization. (This reasoning works for any inner product space; it doesn't have to be a Hilbert space.) We then see that we produce an orthonormal sequence $f_1, f_2, \cdots$ (under the R inner product), such that any function $f \in \mathcal{S}(\mathbb{R}^d)$ is written in the form

$$f = \sum_{n=1}^\infty a_n f_n, \quad a_n = R(f, f_n).$$

It turns out this sequence is Cauchy; it's easy to show that the partial sums $\sum_{n=1}^N a_n^2$ are bounded by the R-norm $R(f,f)$, and therefore the infinite sum makes sense as a limit $\lim_{N\to\infty} \sum_{n=1}^N a_n f_n$ in the R-topology. (The space is not complete – there will be sums like this where $\sum a_n^2$ is finite but the limit is not a Schwartz function – instead, that would be the completion of our space and that's a Hilbert space.) The exact choice of the basis is not so important, but it is necessary for what comes next:

Definition 32

Let $X_1, X_2, \cdots$ be a sequence of iid standard normal random variables defined on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Then for any Schwartz function $f \in \mathcal{S}(\mathbb{R}^d)$, which we write as $f = \sum_{n=1}^{\infty} a_n f_n$, we may define the random variable

$$\phi(f) = \sum_{n=1}^{\infty} a_n X_n;$$

this infinite series converges in $L^2(\Omega)$. Then ϕ is the Euclidean free field (for mass m).

This is similar to how we construct Ito integrals; we can't do so pointwise but we can make a series of approximations which converge in L^2. Notice that this gives us the covariance structure for two functions $f = \sum a_n f_n, g = \sum b_n f_n$

$$\mathrm{Cov}(\phi(f), \phi(g)) = \mathrm{Cov}\left(\sum a_n X_n, \sum b_n X_n\right) = \sum a_n b_n = R(f, g),$$

which is exactly what we wanted. Heuristically, we can interpret $\phi(f)$ as $\int f(x)\phi(x)dx$, even though this doesn't make sense rigorously because ϕ is not defined pointwise. (Sometimes we want to think of ϕ as a "function of all functions at once," but the tricky thing is establishing continuity if we think of ϕ as living on some abstract space.)

Remark 33. *Usually when people work with the free field (for example in Liouville quantum gravity), we're taking integrals of the free field or exponentials of it along certain paths. In those cases we don't really need ϕ to continually vary, but there might be situations where it matters. (If $f_n \to f$, then it's clear that $\phi(f_n) \to \phi(f)$ in L^2, but we may want to make a statement like "with probability one, $f \mapsto \phi(f)$ is continuous on Schwartz functions for all sequences at once," much like for Brownian motion.)*

Regardless, what we're going to understand now is what we can do with the Euclidean free field. This is abstractly a construction of a random field on $\mathbb{R}^d$, and heuristically we can think of it as "having a probability density function proportional to $\exp\left(-\frac{1}{2}\int_{\mathbb{R}^d} |\nabla\phi|^2 - \frac{m^2}{2}\int \phi^2\right)$" with respect to "Lebesgue measure on the space of fields." The question, though, is why this will help us construct this path integral or Hamiltonian that we're after. That hasn't been made clear yet from last week's lectures, especially since we approached the Schrodinger equation case very differently (there we integrated from 0 to t, but here we integrate over all time). What we'll present now is a **probabilistic interpretation of the Osterwalder-Schrader axioms** for formalizing quantum field theories and where they come from.

For this, we'll need to go back to a single-particle system. Take $\mathcal{H} = L^2(\mathbb{R})$, and let S be our action; assume that it has the **additive property**, meaning that for two paths $\gamma_1 : [t_0, t_1] \to \mathbb{R}$ and $\gamma_2 : [t_1, t_2] \to \mathbb{R}$ with $\gamma_1(t_1) = \gamma_2(t_1)$, we can let $\gamma : [0, t_2] \to \mathbb{R}$ be the concatenation of those two paths and we have

$$S(\gamma) = S(\gamma_1) + S(\gamma_2).$$

This has been true for all of the actions we've seen so far. Then if we have the path integral form

$$U(t)\psi(x) = \frac{1}{Z(t)} \int_{\substack{\gamma:[0,t]\to\mathbb{R} \\ \gamma(0)=x}} \psi(\gamma(t)) e^{iS(\gamma)} \mathcal{D}\gamma,$$

the typical **probabilistic approach** is to perform Wick rotation to get the Euclidean action S_E and Euclidean normalizing constant Z_E and then try to get a rigorous meaning of the semigroup

$$P(t)\psi(x) = \frac{1}{Z_E(t)} \int_{\substack{\gamma:[0,t]\to\mathbb{R} \\ \gamma(0)=x}} \psi(\gamma(t)) e^{-S_E(\gamma)} \mathcal{D}\gamma,$$

proving in particular that $P(t)$ has the necessary properties. But in general it's difficult to figure out Z_E (that is, write down the right normalizing constant that lets us write this as an expectation without having dependence on x), because there was a problem with the potential V and now we have an even bigger problem when we have fields.

So the **alternative approach** is as follows: take that same Hilbert space $\mathcal{H} = L^2(\mathbb{R})$, and now given any positive measure μ on $\mathbb{R}$ **with a density** we may consider the Hilbert space $\tilde{\mathcal{H}} = L^2(\mu)$ instead. If a particle is in state $\tilde{\psi} \in \tilde{H}$, then we interpret its probability density **with respect to** μ as being proportional to $|\tilde{\psi}|^2$, just like with Lebesgue measure. There's no loss in moving to this new system, since a state $\psi \in \mathcal{H}$ is equivalent to a state $\tilde{\psi} \in \tilde{\mathcal{H}}$ via $\tilde{\psi}(x) = \frac{\psi(x)}{\sqrt{f(x)}}$:

$$\int_A |\psi(x)|^2 dx = \int_A |\tilde{\psi}(x)|^2 f(x) dx = \int_A |\tilde{\psi}(x)|^2 d\mu(x).$$

When we work with this other measure, the evolution group $\tilde{U}(t)$ should be such that the evolution is exactly the same as U, meaning that if $\tilde{\psi}_0 = \frac{\psi_0}{\sqrt{f}}$ and $\tilde{\psi}_t = \tilde{U}(t)\psi_0$, $\psi_t = U(t)\psi_0$, then we should also have $\tilde{\psi}_t = \frac{\psi_t}{\sqrt{f}}$. We thus have

$$\begin{aligned}
\tilde{U}(t)\tilde{\psi}_0(x) = \tilde{\psi}_t(x) &= \frac{\psi_t(x)}{\sqrt{f(x)}} \\
&= \frac{1}{Z(t)\sqrt{f(x)}} \int_{\gamma:\gamma(0)=x} \psi_0(\gamma(t)) e^{iS(\gamma)} \mathcal{D}\gamma \\
&= \frac{1}{Z(t)\sqrt{f(x)}} \int_{\gamma:\gamma(0)=x} \sqrt{f(\gamma(t))} \tilde{\psi}_0(\gamma(t)) e^{iS(\gamma)} \mathcal{D}\gamma,
\end{aligned}$$

where the blue parts are the parts different from the Lebesgue measure case. Similarly, in the Euclidean version, we should have

$$\tilde{P}(t)\tilde{\psi}_0(x) = \frac{1}{Z_E(t)\sqrt{f(x)}} \int_{\gamma:\gamma(0)=x} \sqrt{f(\gamma(t))} \tilde{\psi}_0(\gamma(t)) e^{-S_E(\gamma)} \mathcal{D}\gamma.$$

The point now is that **if we can construct** a random path $\gamma : \mathbb{R} \to \mathbb{R}$ over all time whose probability density is proportional to $e^{-S_E(\gamma)}$ (in other words, if we can construct such a time-indexed stochastic process), and γ is (1) invariant under time-translation, (2) invariant under time-inversion, and (3) Markovian, then γ is a stationary Markov process and we can let μ be the law of $\gamma(0)$ (a typical point). Then we can rigorously construct $\tilde{P}(t)$ by taking $Z_E(t) = 1$ identically and having

$$P(t)\tilde{\psi}(x) = \mathbb{E}\left[\tilde{\psi}(\gamma(t)) | \gamma(0) = x\right]$$

defined as a conditional expectation, and by the Markov property this will be a semigroup.

So our goal in the particle case is to construct a path with certain properties, and this framework also carries over to fields or more general objects as well. (Time inversion and being Markovian is equivalent to what the axioms call "reflection positivity.")

Remark 34. *If this construction works, then it turns out we should have also had the original Z_E equal to 1. This might seem confusing, but the idea is that with the Schrodinger system (a single particle in a potential), we only did things in the case where V is bounded below. But this $e^{-S_E(\gamma)}$ thing can be done only when V is growing, since otherwise we can't have a stationary path with a density. So this construction is more restrictive in that sense (for example, we can't do this program with a free particle), but it's something different.*

And sometimes we have to modify S via renormalization terms to make this possible – for example in the case of an x^2 potential, we have to put a $-\frac{1}{2}$ term as well. That explains why in our one-dimensional particle case earlier, we did not have $Z_E = 1$. (This will be elaborated on the next lecture!)

10 January 29, 2025

We'll continue our construction of semigroups via stochastic processes today. First, we'll elaborate on the point about the normalization constant from last time.

Recall that if S_E has a certain additive structure, then we formally want to construct a measure on paths $\gamma : \mathbb{R} \to \mathbb{R}$ with density $e^{-S_E(\gamma)}$; such a probability measure is typically constructed as a limit of measures on paths $\gamma : [-T, T] \to \mathbb{R}$ as $T \to \infty$, and the measure has density $\frac{1}{C(T)} e^{-S_E(\gamma)}$ with normalizing constant

$$C(T) = \int_{\gamma:[-T,T]\to\mathbb{R}} e^{-S_E(\gamma)} \mathcal{D}\gamma.$$

But the additive structure of S_E implies that $C(T+T') = C(T)C(T')$, so in fact $C(T) = e^{kT}$ for some fixed constant k. Thus defining the modified Euclidean action

$$S'_E(\gamma) = S_E(\gamma) + k(T_1 - T_0)$$

for any path $\gamma : [T_0, T_1] \to \mathbb{R}$ (in other words, we're adding a constant to the Lagrangian density), the density is exactly $e^{-S'_E(\gamma)}$ with no additional proportionality constant. That's what we're doing when we treat $Z_E(t) = 1$ in our framework, and it also shows (at least formally) that the density of paths on the full real line $\gamma : \mathbb{R} \to \mathbb{R}$ also has density $e^{-S'_E(\gamma)}$.

Turning back to what we had before, we're trying to consider

$$\tilde{P}(t)\tilde{\psi}(x) = \frac{1}{Z_E(t)\sqrt{f(x)}} \int_{\substack{\gamma:[0,t]\to\mathbb{R} \\ \gamma(0)=x}} \tilde{\psi}(\gamma(t)) e^{-S_E(\gamma)} \sqrt{f(\gamma(t))} \mathcal{D}\gamma,$$

where f is chosen to be the probability density function (pdf) of $X(0)$ for X a path chosen from the measure ν on random paths. (Remember that the distribution of $X(t)$ is the same for any t by translation-invariance.)

Proposition 35

(This is a heuristic claim rather than a rigorous one.) Define the operators $Q(t)$ via

$$Q(t)\tilde{\psi}(x) = \mathbb{E}\left[\tilde{\psi}(X(t)) \middle| X(0) = x\right].$$

If X is a stationary Markov process (meaning that the distribution is invariant under time-translation) and the law of X is invariant under time inversion, then $Q(t) = \tilde{P}(t)$ and $Z_E(t) = e^{kt}$ for the constant k above.

Once we've accepted this justification, the rest will be rigorous: we'll define a Markov process which satisfies these conditions, and we'll prove that Q has the necessary properties (strong continuity and so on) so that we can use Hille-Yosida. The advantage of such a rigorous construction is that unlike in the quantum framework, we can actually extract information without going into exact formulas (understanding the nature of the system by qualitative analysis that analysts do).

Proof. First, note that

$$\mathbb{P}(|X(0) - x| < \varepsilon) = f(x) \cdot 2\varepsilon + o(\varepsilon)$$

because f is the density of $X(0)$, but at the same time we also have

$$\begin{aligned}\mathbb{P}(|X(0)-x|<\varepsilon) &= \int_{\substack{\gamma:\mathbb{R}\to\mathbb{R}\\ |\gamma(0)-x|<\varepsilon}} e^{-S_E'(\gamma)}\mathcal{D}\gamma \\ &= \left(\int_{\substack{\gamma:(-\infty,0]\to\mathbb{R}\\ |\gamma(0)-x|<\varepsilon}} e^{-S_E'(\gamma)}\mathcal{D}\gamma\right)\left(\int_{\substack{\gamma:[0,\infty)\to\mathbb{R}\\ |\gamma(0)-x|<\varepsilon}} e^{-S_E'(\gamma)}\mathcal{D}\gamma\right)\end{aligned}$$

by breaking up the paths into the part before and after 0 and using the additive nature of S_E (and hence S_E'). But by time-inversion these two terms are equal, so this last expression is $\left(\int_{\substack{\gamma:(-\infty,0]\to\mathbb{R}\\ |\gamma(0)-x|<\varepsilon}} e^{-S_E'(\gamma)}\mathcal{D}\gamma\right)^2$, meaning that (setting our expressions for the probability equal) we have

$$\int_{\substack{\gamma:(-\infty,0]\to\mathbb{R}\\ |\gamma(0)-x|<\varepsilon}} e^{-S_E'(\gamma)}\mathcal{D}\gamma = \sqrt{f(x)2\varepsilon+o(\varepsilon)}.$$

Therefore, going back to our definition of Q, we can write the conditional expectation as

$$\begin{aligned}Q(t)\tilde{\psi}(x) &= \mathbb{E}\left[\tilde{\psi}(X(t))\middle|X(0)=x\right] \\ &= \lim_{\varepsilon\to 0}\frac{\int_{\gamma:\mathbb{R}\to\mathbb{R},|\gamma(0)-x|<\varepsilon}\tilde{\psi}(\gamma(t))e^{-S_E'(\gamma)}\mathcal{D}\gamma}{\mathbb{P}(|X(0)-x|<\varepsilon)}.\end{aligned}$$

We can now break up this numerator into three parts (fixing the value of $\gamma(t)=y$, then integrating over y) and plug in the denominator estimate from above, so that we get (note here that we're kind of thinking of $\mathcal{D}\gamma$ as integrating over $d\gamma(s)$ for all $s\in\mathbb{R}$, and so we're taking the one at t and making it our "dy")

$$\lim_{\varepsilon\to 0}\frac{\int\left(\int_{\substack{\gamma:(-\infty,0]\to\mathbb{R}\\ |\gamma(0)-x|<\varepsilon}} e^{-S_E'(\gamma)}\mathcal{D}\gamma\int_{\substack{\gamma:[0,t]\to\mathbb{R}\\ |\gamma(0)-x|<\varepsilon,|\gamma(t)-y|<\varepsilon}} e^{-S_E'(\gamma)}\mathcal{D}\gamma\int_{\substack{\gamma:[t,\infty)\to\mathbb{R}\\ |\gamma(t)-y|<\varepsilon}} e^{-S_E'(\gamma)}\mathcal{D}\gamma\right)\tilde{\psi}(y)dy}{f(x)2\varepsilon+o(\varepsilon)}.$$

Now as $\varepsilon\to 0$, the first and third integrals are $\sqrt{f(x)2\varepsilon+o(\varepsilon)}$ and $\sqrt{f(y)2\varepsilon+o(\varepsilon)}$ by translation and time-inversion invariance. Thus this simplifies to

$$\begin{aligned}Q(t)\tilde{\psi}(x) &= \lim_{\varepsilon\to 0}\frac{\int\sqrt{f(x)2\varepsilon+o(\varepsilon)}\sqrt{f(y)2\varepsilon+o(\varepsilon)}\left(\int_{\substack{\gamma:[0,t]\to\mathbb{R}\\ |\gamma(0)-x|<\varepsilon,|\gamma(t)-y|<\varepsilon}} e^{-S_E'(\gamma)}\mathcal{D}\gamma\right)\tilde{\psi}(y)dy}{f(x)2\varepsilon+o(\varepsilon)} \\ &= \frac{1}{e^{kt}\sqrt{f(x)}}\int_{\gamma:[0,t]\to\mathbb{R},\gamma(0)=x} e^{-S_E(\gamma)}\tilde{\psi}(\gamma(t))\mathcal{D}\gamma,\end{aligned}$$

where in the last equality we perform the integral over y (using that $\sqrt{f(y)2\varepsilon}$ corresponds to the density at x for any time t) and use that $S_E'(\gamma)=S_E(\gamma)+kt$. This is exactly our desired expression. □

Remark 36. *The interpretation of k is as the "smallest eigenvalue of the Hamiltonian;" note that we didn't get this constant in the Feynman-Kac formula before because we were working on $L^2(\mathbb{R})$ and now we're working on $L^2(\nu)$ so that we have a stationary process.*

Remark 37. *Remember that all of this is not rigorous (we haven't defined the measure on paths or what space we're working on in general, for example). In fact, the reasoning above doesn't work for Brownian motion – we can't generate a time-invariant Brownian motion because we need a density proportional to* $\exp\left(-\frac{1}{2}\int_{-\infty}^{\infty}\dot{X}(t)^2dt\right)$*. But we can in fact generate something like a process $(X(t))_{t\in\mathbb{R}}$ which has density proportional to* $\exp\left(-\frac{1}{2}\int_{-\infty}^{\infty}\dot{X}(t)^2dt-\frac{m^2}{2}\int_{-\infty}^{\infty}X(t)^2dt\right)$ *(once we make sense of it). In general the point is that we do need a growing potential; from the covariance structure*

we can check that the covariance structure $\mathrm{Cov}(X(s), X(t))$ *matches the Ornstein-Uhlenbeck process. And the physicists basically know what is right and what is not through discourse (looking at formulas and then getting evidence); once something is well-established, there's a good chance it can be made mathematically rigorous.*

Turning back to rigor now, we're interested in a quantum field $\phi(t, x)$ (where $\phi : \mathbb{R}^d \to \mathbb{R}$ for $d = 4$, and for each fixed time t we get a map $\phi(t, \dot{c}) : \mathbb{R}^{d-1} \to \mathbb{R}$ which is the fundamental object we're trying to construct a quantum system for). Following this prescription, this map $\mathbb{R}^d \to \mathbb{R}$ is the "time trajectory of our system" (which takes the place of what was previously $\gamma : \mathbb{R} \to \mathbb{R}$). We constructed a random field that formally has density proportional to $\exp\left(-\frac{1}{2}\int |\nabla\phi|^2 - \frac{m^2}{2}\int \phi^2\right)$, but our next task is to understand ϕ as a time-indexed stochastic process much like $X(t)$.

We'll soon show that ϕ is Markovian and translation and time-inversion invariant, but right now notice that we don't even really understand how to take a time-slice of ϕ – all we know is that ϕ is defined as the collection of random variables $\{\phi(f) : f \in \mathcal{S}(\mathbb{R}^d)\}$, where the $\phi(f)$ are jointly centered Gaussian random variables with covariance

$$\mathrm{Cov}(\phi(f), \phi(g)) = R(f, g) = \int f(x)g(y)G_m(x - y)dxdy.$$

So the question is how to define this at time t – we would want to integrate against functions that are supported only at time t (so they're like a function in $\mathbb{R}^{d-1}$ times the Dirac delta at t). Those are obviously not Schwartz functions, but it turns out they are in the **closure** of these Schwartz functions under the R-topology (which is a Hilbert space) and thus we can define them. What we'll do is take $f \in \mathcal{S}(\mathbb{R}^{d-1})$ and show that $\tilde{f}(s, x) = f(x)\delta_t(s)$ is in the completion so that $\phi(\tilde{f})$ makes sense.

Remark 38. *There's also a discrete construction of the Gaussian free field, and we can move from discrete to continuous and prove limit theorems. But we won't get into that much here.*

11 January 31, 2025

We've previously constructed the Euclidean free field ϕ with mass $m \geq 0$ on $\mathbb{R}^d$, and we want to think of this field as a trajectory of fields on $\mathbb{R}^{d-1}$ over time. Fortunately, $\phi(t, \cdot)$ does indeed make sense as a Gaussian field. To show this, for any Schwartz function $f \in \mathcal{S}(\mathbb{R}^{d-1})$, we want to make sense of $\int_{\mathbb{R}^{d-1}} \phi(t, x)f(x)dx$. Define $\tilde{f} : \mathbb{R}^d \to \mathbb{R}$ via

$$\tilde{f}(t, x) = f(x)\delta_t(s);$$

this is a distribution rather than a function and clearly not Schwartz, but we'll show that $\tilde{f}$ is in the completion of $\mathcal{S}(\mathbb{R}^d)$ under the R-topology and hence we may define $\phi(\tilde{f})$ via a limit of some Cauchy sequence of Schwartz functions, since f_n being Cauchy means $\phi(f_n)$ is a Cauchy sequence of random variables in L^2 (since ϕ is an isometry).

Fact 39

Note that even though all Schwartz functions are in L^2, the completion doesn't have to be contained in L^2 because we're looking under the R-topology, which integrates against the Green's function. In mathematical terms, ϕ is really a random element of a negative-index Sobolev space.

Formally, the covariance structure of $\phi_t(f) = \int_{\mathbb{R}^{d-1}} \phi(t, x)f(x)dx$ should satisfy (for any $f, g \in \mathcal{S}(\mathbb{R}^{d-1})$)

$$\mathrm{Cov}(\phi_t(f), \phi_s(g)) = \iint_{\mathbb{R}^{d-1}\times\mathbb{R}^{d-1}} f(x)g(y)G_m(t - s, x - y)dxdy$$

(here using that $\text{Cov}(\phi(t,x),\phi(s,y)) = G_m(t-s,x-y)$). But this integral is finite because we've shown that $|G_m(t-s,x-y) \leq \frac{1}{||x-y||^{d-2}}$ and thus if we switch to polar coordinates we're still integrable over $\mathbb{R}^{d-1}$ because we get no singularity at zero. And so we should expect that these random variables $\phi_t(f)$ make sense and have finite variance.

Let's do the proof rigorously now; without loss of generality assume $t = 0$. to approximate our delta functions, let $w : \mathbb{R} \to [0,\infty)$ be any Schwartz function with $\int_{-\infty}^{\infty} w(x)dx = 1$ and define

$$f_n(s,x) = nw(ns)(x).$$

As n increases, $nw(n(s-t))$ becomes a sharper and sharper probability density around t, and we claim that $\{f_n\}_{n\geq 1}$ is Cauchy in the R-topology. To prove this, notice that

$$\begin{aligned}||f_k - f_n||_R &= R(f_k - f_n, f_k - f_n)\\ &= \iiiint (f_k(t,x) - f_n(t,x))(f_k(s,y) - f_n(s,y))G_m(t-s,x-y)dxdysdt\\ &= \iiiint f(x)f(y)\left(kw(kt) - nw(nt)\right)\left(kw(ks) - nw(ns)\right)G_m(t-s,x-y)dxdydsdt.\end{aligned}$$

We wish to show this goes to zero as k, n go to infinity simultaneously. Indeed, when we expand out the parentheses, we will get four separate terms; it suffices to show that all four terms converge to the same limit. Indeed, for example the first term can be rewritten

$$\begin{aligned}&\iiiint f(x)f(y)kw(kt)kw(ks)dsdtdxdyG_m(t-s,x-y)\\ =&\iiint f(x)f(y)w(t')w(s')G\left(\frac{t'}{k} - \frac{s'}{k}, x-y\right)dt'ds'dxdy;\end{aligned}$$

now as $k, n \to \infty$ simultaneously the first argument of G goes to zero, and so by the dominated convergence theorem (since we have good bounds at the origin) this converges to $\iiiint f(x)f(y)w(s')w(t')G_m(0,x-y)dxydt'ds'$, and now we can integrate out s' and t' because we have separated variables and so we just have $\iint f(x)f(y)G_m(0,x-y)dxdy$. The same occurs for the other terms, so we've indeed shown that $\{f_n\}$ is Cauchy, as desired.

By the same argument, we can prove that if we take any other $\tilde{w}$ instead of w, the new sequence we get is equivalent to the old one (the limit doesn't depend on the choice). Thus this is indeed well-defined.

The point is that we can view the whole thing we've constructed as a single object rather than a bunch of random variables:

Definition 40

Let $\mathcal{A}$ be the set of all functions $\psi : \mathcal{S}(\mathbb{R}^{d-1}) \to \mathbb{R}$ (with no restrictions), and for any $f \in \mathcal{S}(\mathbb{R}^{d-1})$ let $E_f : \mathcal{A} \to \mathbb{R}$ be the map

$$E_f(\psi) = \psi(f).$$

In other words, we view ψ as a field on the space of Schwartz functions and evaluate it at a certain point via this evaluation map E_f. Now let $\mathcal{C}$ be the (smallest) σ-algebra generated by these E_fs over all Schwartz functions f. We call $\mathcal{C}$ the **cylinder σ-algebra**.

(A σ-algebra is a collection of subsets containing the empty set and also containing complements and countable unions; **the more formal definition will come next lecture**.) If we now take this collection of random variables $\{\phi_0(f) : f \in \mathcal{S}(\mathbb{R}^{-1})$ that we've already defined, then we can put them together by defining a map $\phi_0 : \Omega \to \mathcal{A}$ (this

is now an $\mathcal{A}$-valued random object) via

$$\phi_0(\omega)(f) = \phi_0(f)(\omega).$$

So at a sample point in our probability space, we can take any element of $\mathcal{A}$ (a map $\mathcal{S}(\mathbb{R}^{d-1}) \to \mathbb{R}$) and evaluate it at our sample point. (This is important formalism because we want to define the Hilbert space; it will turn out to be $L^2(\mathcal{A})$ with the σ-algebra $\mathcal{S}$, together with a measure, which is the law of the $\mathcal{A}$-valued random variable that we're introducing here.)

Lemma 41

ϕ_0 is indeed an $\mathcal{A}$-valued random variable (that is, it is measurable). In other words, for all $C \in \mathcal{C}$, $\phi_0^{-1}(C)$ is in $\mathcal{F}$ for our probability space $(\Omega, \mathcal{F}, \mathbb{P})$.

Proof. The set of all $C \subseteq \mathcal{A}$ that do satisfy $\phi_0^{-1}(C) \in \mathcal{F}$ is indeed a σ-algebra; thus, it suffices to prove the claim just for some Cs that generate $\mathcal{C}$ as a σ-algebra. An example of such a generating set is

$$\{E_f^{-1}(B) : f \in \mathcal{S}(\mathbb{R}^{d-1}), B \in \mathcal{B}(\mathbb{R})\},$$

and so it suffices to show that $\phi_0^{-1}(E_f^{-1}(B)) \in \mathcal{F}$. But by definition we have

$$\begin{aligned}\phi_0^{-1}(E_f^{-1}(B)) &= \left\{\omega : \phi_0(\omega) \in E_f^{-1}(B)\right\} \\ &= \{\omega : \phi_0(\omega)(f) \in B\} \\ &= \{\omega : \phi_0(f)(\omega) \in B\},\end{aligned}$$

and since $\phi_0(f)$ is a random variable (we have already constructed it as a limit of finite sums and thus it is a measurable map) this is indeed in $\mathcal{F}$, completing the proof. □

(Of course, we're doing the timeslice at $t = 0$ here, but we can do this for any t.) Next time, we'll construct our Hilbert space explicitly, and then we'll be in position to prove all the properties we need.

12 February 3, 2025

Last time, we were constructing the free field as a Markov process (rather than just as an abstract Gaussian field on $\mathbb{R}^d$) by looking at time-slices evolving over time. Specifically, we had the space $\mathcal{A}$ of all functions $\mathcal{S}(\mathbb{R}^{d-1}) \to \mathbb{R}$ (inputting a Schwartz function and outputting a real number) and let $\mathcal{C}$ be the cylinder σ-algebera on $\mathcal{A}$ (that is, the smallest σ-algebra of subsets of $\mathcal{A}$ under which the evaluation maps at all Schwartz functions are measurable). Then $\{\phi_t\}_{T \in \mathbb{R}}$ is an $\mathcal{A}$-valued stochastic process defined on some background probability space $(\Omega, \mathcal{F}, \mathbb{P})$ on which we have an iid sequence of standard normals – each ϕ_t is then basically a random distribution, and it's the one we'll use to construct our quantum system.

We'll prove shortly that this is stationary (in the sense that if we shift time, the joint distribution of the collection remains invariant), so in particular each ϕ_t has the same distribution. If we let μ be the law of ϕ_0, meaning that for all cylinder sets $C \in \mathcal{C}$ we have

$$\mu(C) = \mathbb{P}(\phi_0 \in C) = \mathbb{P}\left(\{\omega : \phi_0(\omega) \in C\}\right),$$

then $(\mathcal{A}, \mathcal{C}, \mu)$ is a probability space, and we'll **define our Hilbert space** to be the L^2 space of this:

$$\begin{aligned}\mathcal{H} &= L^2(\mathcal{A}, \mathcal{C}, \mu) \\ &= \left\{ F : F \text{ is a measurable map } \mathcal{A} \to \mathbb{C} \text{ such that } \int_{\mathcal{A}} |F(\phi)|^2 d\mu(\phi) < \infty \right\}.\end{aligned}$$

Before we make the properties of this space clear, we'll first mention that our semigroup will be the map $P_t : \mathcal{H} \to \mathcal{H}$ such that

$$P_t F(\phi) = \mathbb{E}\left[F(\phi_t) | \phi_0 = \phi\right].$$

This will turn out to be a contraction semigroup (which is very helpful for proving strong continuity), and by Hille-Yosida it can be written in the form e^{-tH} for some generator H; that will be the Hamiltonian for the free quantum field and thus we will define the latter as $U(t) = e^{-itH}$ by Stone's theorem. The hardest part will be proving that we do actually have a semigroup (equivalently that ϕ_t is a Markov process), and the way we wrote things down we heuristically kind of have something of the form $\prod_{i=-\infty}^{\infty} P(x_i, x_{i+1})$ (since only the squared gradient appears), but proving it will be some more work.

First, we should clarify what this conditional expectation actually means in the more abstract probabilistic context.

Definition 42

Let Ω be a set and $\mathcal{F}$ a ***σ*-algebra** on Ω, meaning that (1) the emptyset $\varnothing$ is in $\mathcal{F}$, (2) the complement A^c of any set $A \in \mathcal{F}$ is also in $\mathcal{F}$, and (3) if $A_1, A_2, \cdots$ are in $\mathcal{F}$, then $\bigcup_{i=1}^{\infty} A_i$ is also in $\mathcal{F}$. We say that $\mathbb{P}$ is a **probability measure** on $(\Omega, \mathcal{F})$ if $P : \mathcal{F} \to [0, 1]$ is a map with $\mathbb{P}(\Omega) = 1$ and $\sum_{i=1}^{\infty} \mathbb{P}(A_i) = \mathbb{P}\left(\bigcup_{i=1}^{\infty} A_i\right)$ for **disjoint** sets $A_1, A_2, \cdots, \in \mathcal{F}$.

(This last property is important because Lebesgue was trying to generalize the Riemann integral that still works on functions like $1\{x \in \mathbb{Q}\}$; he figured out you need countable additivity rather than finite additivity. There is a literature on probability with finite additivity, but it hasn't gone very far.)

Definition 43

A **random variable** is then a **measurable map** $X : \Omega \to \mathbb{R}$, meaning that for any $A \in \mathcal{B}(\mathbb{R})$ in the **Borel *σ*-algebra** (the smallest σ-algebra containing all open sets), $X^{-1}(A) = \{\omega : X(\omega) \in \mathcal{A}\}$ is an element of $\mathcal{F}$. More generally for any measurable space $(S, \mathcal{S})$, an ***S*-valued random variable** is a measurable map $X : \Omega \to S$, meaning that $X^{-1}(A) \in \mathcal{F}$ for all $A \in \mathcal{S}$.

For any random variable X, the ***σ*-algebra generated by *X***, denoted $\sigma(X)$ is the smallest σ-algebra $\mathcal{G}$ such that X is measurable with respect to $\mathcal{G}$. (The intersection of σ-algebras is always a σ-algebra, so we get the "smallest" one by taking the intersection of all such objects.)

Note that the Borel σ-algebra is very huge – it includes things like the Cantor set and basically any other object we'd typically encounter. And we can think of the σ-algebra generated by X as all events determined by X (so things of the form "X is in some set A"), since explicitly we have $\sigma(X) = \{X^{-1}(A) : A \in \mathcal{S}\}$.

Fact 44

Let $X = \{X_i\}_{i \in I}$ be any arbitrary collection of real-valued random variables. We may view X as an S-valued random variable for $S = \prod_{i \in I} \mathbb{R}$, the set of all maps from I into $\mathbb{R}$, where $\mathcal{S}$ is the cylinder σ-algebra (the smallest σ-algebra such that the maps $f \mapsto f(i)$ are measurable for all $i \in I$). This is the product σ-algebra if I is countable but something else otherwise; this is kind of like the product topology.

In other words, collections of random variables may always be viewed as single random variables.

Lemma 45

Let X be an S-valued random variable defined on Ω (for some measurable space S; we're suppressing the σ-algebras here). Suppose $Y : \Omega \to \mathbb{R}$ is measurable with respect to $\sigma(X)$ (that is, we have a random variable which is $\sigma(X)$-measurable). Then there is some measurable function $F : S \to \mathbb{R}$ such that $Y = F(X) = F \circ X$ with probability 1 (and if there is another one F', then $\mathbb{P}(F(X) = F'(X)) = 1$).

(This can perhaps be thought of as a kind of "universal property.")

Proof. One important thing is that unless we use transfinite induction, we cannot use induction in the typical way when dealing with Borel sets because it's difficult to build up sets from open sets. Instead, the way we typically do measure-theoretic proofs is to **first show the property when $Y = 1_A$ is an indicator function** (meaning that it takes value 1 if $\omega \in A$ and 0 otherwise). Indeed, $1_A \in \sigma(X)$ implies that $A = X^{-1}(C)$ for some $C \in \mathcal{S}$, which means that $Y(\omega) = 1$ if $X(\omega) \in C$ and $Y(\omega) = 0$ otherwise. We can then define $F : S \to \mathbb{R}$ to be the indicator function 1_C and clearly we have $Y = F \circ X$ and $F : S \to \mathbb{R}$ is indeed measurable.

Next, we consider the case of **simple functions**, meaning that $Y = \sum_{i=1}^n c_i 1_{A_i}$ for some integer n, some $c_i \in \mathbb{R}$, and some $A_i \in \sigma(X)$. For this case, we can find F_i such that $1_{A_i} = F_i \circ X$ and let $F = \sum_{i=1}^n c_i F_i$; indeed we can check that $Y = F \circ X$.

The last step uses the following lemma (which we won't prove):

Lemma 46

Let $\mathcal{G}$ be a σ-algebra on Ω and let $g : \Omega \to \mathbb{R}$ be measurable. Then there exists a sequence $\{g_n\}$ of measurable simple functions (with respect to $\mathcal{G}$) such that $g_n \to g$ pointwise.

It can be verified that measurability of real-valued functions (or functions taking values in any Polish space) is preserved under pointwise limits (by checking that inf and sup are measurable, meaning that lim inf and lim sup are as well, and if a limit exists then it's equal to the lim sup). Thus for any $\mathcal{G}$-measurable $Y : \Omega \to \mathbb{R}$, there exists a sequence of measurable simple functions Y_n such that $Y_n \to Y$ pointwise. By our previous step, for each n we can find a measurable $F_n : S \to \mathbb{R}$ with $Y_n = F_n \circ X$. This means that for all $\omega \in \Omega$ we have

$$Y(\omega) = \lim Y_n(\omega) = \lim F_n(X(\omega)),$$

meaning that F_n has a pointwise limit **on the image of** X. So now we can take $F = \limsup_n F_n$ (if we just try to take a limit we run into some measurability issues), and we will have $Y = F(X)$ with probability 1 as desired. □

This finally allows us to define conditional expectation; our goal is to define something like $\mathbb{E}[Y|X = x]$, where $X : \Omega \to S$ and $Y : \Omega \to \mathbb{R}$ are measurable. Without putting the assumption that $X = x$, we can define $\mathbb{E}[Y|X]$ to be the random variable satisfying the following properties:

1. it is $\sigma(X)$-measurable (meaning it is a function of X), and
2. for any $\sigma(X)$-measurable random variable Z, we have

$$\mathbb{E}\left[\mathbb{E}[Y|X]Z\right] = \mathbb{E}[YZ]$$

if the right-hand side is well-defined.

The point is that by Lemma 45, $\mathbb{E}[Y|X] = F(X)$ almost surely for some measurable $F : S \to \mathbb{R}$, and we will define $\mathbb{E}[Y|X = x] = F(x)$. (Typically we don't go to this level of specificity when we work abstractly with conditional expectation, but we'll need it for the purposes of our course!) So F will take in a field and map it to a real number, and we'll see that in action next time.

13 February 5, 2025

We previously constructed the free field as a stochastic process on a particular Hilbert space; today, we'll prove that we've constructed a semigroup that satisfies the conditions of the Hille-Yosida theorem. Then we'll move on to a similar but more complicated construction for Maxwell theory (the free electromagnetic field).

Theorem 47

The process $\{\phi_t\}_{t\in\mathbb{R}}$ we have constructed is a **Markov process**. The meaning of this statement is as follows: define the σ-algebras $\mathcal{F}_{\leq t} = \sigma\left(\{\phi_s : s \leq t\}\right)$, $\mathcal{F}_t = \sigma(\phi_t)$, and $\mathcal{F}_{>t} = \sigma\left(\{\phi_s : s > t\}\right)$. Then $\{\phi_t\}$ is **Markov** if for all real-valued random variables X that are $\mathcal{F}_{>t}$-measurable, we have almost surely that

$$\mathbb{E}\left[X|\mathcal{F}_{\leq t}\right] = \mathbb{E}\left[X|\mathcal{F}_t\right].$$

In other words, the information of the random variable given only the present time is as accurate as having the present but also all past times.

We gave the definition of conditional expectation last time – recall that the idea is for any measurable map $X : \Omega \to \mathbb{R}$ with $\mathbb{E}[|X|] < \infty$ (we say that X is **integrable** in such a case), for any $\mathcal{G} \subseteq \mathcal{F}$ we can define $\mathbb{E}[X|\mathcal{G}]$ to be the $\mathcal{G}$-measurable random variable satisfying $\mathbb{E}[XY] = \mathbb{E}[\mathbb{E}[X|\mathcal{G}]Y]$ for all $\mathcal{G}$-measurable Y satisfying $\mathbb{E}[|XY|] < \infty$. (Similarly we define $\mathbb{E}[X|Z] = \mathbb{E}[X|\sigma(Z)]$ for any random variable Z.) It's a bit hard to convey what this means in a short time, but we can think of a σ-algebra as follows.

Example 48

Consider the unit square $[0,1]^2$ as a sample space (so our experiment yields some point in the square). We can consider the Borel σ-algebra of this, but we can also consider a sub-σ-algebra by partitioning the square into smaller squares. The sub-σ-algebra generated by those squares is the set of all unions of squares, so it only captures the information of whether "the point belongs to these particular squares." So if we have a random variable X which is defined on the unit square, the conditional expectation given this sub-σ-algebra is the function constant on each square, and its value is the average on each square. So $\mathbb{E}[X|\mathcal{G}]$ is the "best prediction we can make about X given that we have only information contained in $\mathcal{G}$."

Example 49

Here's another more concrete way to think about this in the case of discrete random variables X, Z: we know that

$$\mathbb{E}[X|Z=k] = \sum_j \mathbb{P}(X=j|Z=k)\cdot j.$$

Now the quantity on the right-hand side is some function $f(k)$ of k, with the property that

$$\begin{aligned}\mathbb{E}[f(Z)g(Z)] &= \sum_k f(k)g(k)\mathbb{P}(Z=k)\\ &= \sum_{k,j} \mathbb{P}(X=j|Z=k)jg(k)\mathbb{P}(Z=k)\\ &= \sum_{k,j} \mathbb{P}(X=j,Z=k)jg(k)\\ &= \mathbb{E}\left[Xg(Z)\right].\end{aligned}$$

This is exactly the characterizing property that we had in our definition, and it's the criterion that's easy to make abstract.

Conditional expectations are a bit weird: phrasing questions may give you different answers (see the **Borel paradox** regarding conditioning on measure-zero events), but the way we set things up gives us the correct resolution.

So going back to the definition of a Markov process in general, the point is that conditional expectations on the past and present only depend on the present. To establish that $\{\phi_t\}$ is indeed Markov, one convenient way to do it is to show that given any t and any Y which is $\mathcal{F}_{>t}$-measurable, we can represent it as $F(X_t, Z)$ for Z **independent** of $\mathcal{F}_{\leq t}$ (meaning that the σ-algebras generated by those two things are independent, which means that $\mathbb{P}(A\cap B) = \mathbb{P}(A)\mathbb{P}(B)$ for any A in one σ-algebra and B in the other). Then we can show that the conditional expectation of such a random variable given $\mathcal{F}_{\leq t}$ is just the same as conditioning on time t:

$$\mathbb{E}\left[F(X_t,Z)|\mathcal{F}_{\leq t}\right] = \mathbb{E}\left[F(X_t,Z)|\mathcal{F}_t\right].$$

(For example, we can think of Brownian motion at some time t_1 as being the value at time t_0 plus whatever happens after time t_0, so that makes Brownian motion a Markov process.)

We need to prove this property for the free field now – in our case we will follow a certain prescription which is not very simple (unlike Brownian motion). For this, fix some $s < t$; we will express $\phi_t(f)$ (where $f \in \mathcal{S}(\mathbb{R}^{d-1})$) as a function of ϕ_s (the **whole field** at time s) and a random field $\psi_{s,t}$ that is independent of $\{\phi_u\}_{u\leq s}$. In fact, this will be additive in the sense that $\psi_{s,t}$ is a function of ϕ_s, and $\phi_t(f) - \psi_{s,t}$ is independent of $\mathcal{F}_{\leq s}$.

The natural guess for $\psi_{s,t}$, since we want to make a Gaussian random variable independent of everything else, is to subtract off the projection in the L^2 space:

$$\psi_{s,t} = \mathbb{E}\left[\phi_t|\phi_s\right].$$

(To make this make sense we basically need to evaluate both sides at some function f.) It turns out this will end up being a smooth version of ϕ_s – this means that as f approaches a delta function the right-hand side will not actually blow up. In fact, $\psi_{s,t}$ will actually be a smooth stationary random field on $\mathbb{R}^{d-1}$, rather than a distribution (and we'll determine what it is). But the formula is not terribly simple: it'll turn out that $\mathbb{E}\left[\phi_t|\phi_s\right]$ is equal to $\mathbb{E}[\phi_t|\phi_s]$ (not for all Gaussian fields, but for this one) and that's what will make it Markovian.

Fact 50

Notice that we're not taking an arbitrary function of the future field, just the field itself; the reason we can do this is because we're dealing with Gaussians. The fact that subtracting off conditional expectations makes us independent of the things we subtracted off is special to Gaussians; typically we only get that they are uncorrelated, which is weaker.

Towards our goal, we can define the kernel (now let $t > 0, x \in \mathbb{R}^{d-1}$ be arbitrary)

$$K_m(t,x) = \int_0^\infty \frac{t}{\sqrt{4\pi s^3}(4\pi s)^{(d-1)/2}} \exp\left(-\frac{t^2 + ||x||^2}{4s} - m^2 s\right) ds.$$

Formally, we then want

$$\psi_{s,t}(x) = \int_{\mathbb{R}^{d-1}} K_m(t-s, x-y)\phi_s(y)dy,$$

and so for an arbitrary function $f \in \mathcal{S}(\mathbb{R}^{d-1})$ **we define**

$$\boxed{\psi_{s,t}(f) = \phi_s(K_m^{t-s} f)},$$

where K_m^t is convolution with the kernel via

$$K_m^t f(x) = \int_{\mathbb{R}^d} K_m(t, x-y) f(y) dy.$$

So basically we "further smooth out f" and then apply ϕ_s to it. (As we will see, this particular thing has some properties so that the random variable $\phi_t(f) - \psi_{s,t}(f)$ becomes uncorrelated, hence independent, of any $\phi_u(g)$. And we can think of this as sort of "being able to solve the Cauchy problem even for distributions.")

To show this, there must be some work involved – this along with time-inversion (easy to show) implies reflection positivity, which is a nontrivial fact. We'll prove two lemmas:

Lemma 51

The Green's function can be expressed as

$$G_m(t,x) = \int_{\mathbb{R}^{d-1}} G_m(0,y) K_m(t, x-y) dy.$$

In other words, $G_m(t)$ is the convolution of $G_m(0)$ and $K_m(t)$.

Lemma 52

K_m behaves nicely under convolution; that is, for all $s, t > 0$,

$$K_m(s+t, x) = \int_{\mathbb{R}^{d-1}} K_m(s,y) K_m(t, x-y).$$

Once we have established those two lemmas, we thus see that

$$G_m(t+s, x) = \int G_m(s,y) K_m(t, x-y) dy,$$

since (letting $*$ denote convolution)

$$G_m^{t+s} = G_m^0 * K_m^{t+s} = G_m^0 * K_m^s * K_m^t = G_m^s * K_m^t$$

by associativity of convolutions. So even though we can't get closure under convolution by G, we still get a relation via K instead. Once we have this, we're done, since for any $u < s < t$ we now want to prove that $\phi_t - \psi_{s,t} = \phi_t - (K_m^{t-s} * \phi_s)$ is independent of ϕ_u. For Gaussian fields we just need to check that covariance is zero when applied to any Schwartz functions f, g, and thus we wish to show that

$$\begin{aligned} &\operatorname{Cov}(\phi_t(f) - \psi_{s,t}(f), \phi_u(g)) \stackrel{?}{=} 0 \\ &\quad \iff \operatorname{Cov}(\phi_t(f), \phi_u(g)) \stackrel{?}{=} \operatorname{Cov}(\psi_{s,t}(f), \psi_u(g)). \end{aligned}$$

We'll just show this for $f = \delta_x$ and $g = \delta_y$ (even though they're not Schwartz functions); indeed, the left-hand side is

$$\operatorname{Cov}(\phi_t(\delta_x), \phi_u(\delta_y)) = \operatorname{Cov}(\phi(t,x), \phi(u,y)) = G_m(t-u, x-y)$$

(we can check this by integrating against smooth functions), while

$$\begin{aligned} \operatorname{Cov}(\psi_{s,t}(\delta_x), \phi_u(\delta_y)) &= \operatorname{Cov}(K_m^{t-s} * \phi_s(\delta_x), \phi_u(\delta_y)) \\ &= \operatorname{Cov}\left(\int K_m(t-s, x-z) \phi_s(z) dz, \phi_u(y) \right) \\ &= \int K_m(t-s, x-z) \operatorname{Cov}(\phi_s(z), \phi_u(y)) dz \\ &= \int K_m(t-s, x-z) G(s-u, z-y) dz \\ &= K_m^{t-s} * G^{s-u}(x-y) \\ &= G_m^{t-u}(x-y), \end{aligned}$$

where the last step is deduced from our lemmas. So we see that the key property we need is that G_m^{s+t} is a convolution of G_m^s with K_m^t, and if we can find a kernel which satisfies this property, we're automatically done – what's left is to check that the K_m^t we wrote down does do so. (Rigorously we need to use functions instead of delta functions in all of this and use double integrals, but the argument basically goes through in the same way.)

We'll close with the intuition for where this K_m^t construction comes from: what's really going on is that we start from time s and start from some point, say the origin. We have some point $(s+t, x)$ that we want to reach, and we can write

$$G_m(s+t, x) = \mathbb{E}\left[\int_0^\infty e^{-m^2 u} \left(\text{pdf of } d\text{-dimensional Brownian motion at time } u \text{ and location } (s+t, x) \right) du \right],$$

where the $e^{-m^2 u}$ can be interpreted as having an $m^2 \delta t$ chance of being killed at any particular time interval δt. We can decompose this expectation by conditioning on the time-coordinate hitting s starting from zero; the expected time for first passage time of Brownian motion gives us the $\frac{1}{\sqrt{s^3}}$ factor, and we'll see the rest of the details next time.

14 February 7, 2025

Last time, we wrote out an explicit kernel $K_m(t, x)$, which takes the form

$$K_m(t,x) = \int_0^\infty \frac{t}{\sqrt{4\pi s^3} (4\pi s)^{(d-1)/2}} \exp\left(-\frac{t^2 + ||x||^2}{4s} - m^2 s \right) ds.$$

In particular, this looks quite similar to the Green's function

$$G_m(t,x) = \int_0^\infty \frac{1}{(4\pi s)^{d/2}} \exp\left(-\frac{t^2+||x||^2}{4s} - m^2 s\right).$$

Last time, we stated Lemma 51 and Lemma 51 as the final steps needed to finish our proof that ϕ_t is Markov. We'll prove these lemmas now:

Proof of Lemma 51. We can write out the expression

$$\begin{aligned} G_m^0 * K_m^t &= \int_{\mathbb{R}^{d-1}} G_m(0,y) K_m(t, x-y)\, dy \\ &= \int_{\mathbb{R}^{d-1}} \int_0^\infty \int_0^\infty \frac{t}{(4\pi u)^{d/2}\sqrt{4\pi s^3}(4\pi s)^{(d-1)/2}} \exp\left(-\frac{||y||^2}{4u} - \frac{t^2+||x-y||^2}{4s} - m^2(s+u)\right) ds du dy. \end{aligned}$$

We'll carry out the outer integration over $\mathbb{R}^{d-1}$ first. We have that

$$\int_{\mathbb{R}^{d-1}} \frac{1}{(4\pi s)^{(d-1)/2}(4\pi u)^{(d-1)/2}} \exp\left(-\frac{||y||^2}{4u} - \frac{||x-y||^2}{4s}\right) dy$$

is just a convolution of Gaussian densities on $\mathbb{R}^{d-1}$ with variance parameters $2u$ and $2s$, respectively, and thus this evaluates to just $\frac{1}{(4\pi(s+u))^{(d-1)/2}} \exp\left(-\frac{||x||^2}{4(s+u)}\right)$. Plugging this back into our expression above,

$$G_m^0 * K_m^t = \int_0^\infty \int_0^\infty \frac{t}{\sqrt{4\pi u}\sqrt{4\pi s^3}} \exp\left(-\frac{t^2}{2s} - \frac{||x||^2}{4(s+u)} - m^2(s+u)\right) ds du.$$

Under the change of variables $r = s+u$, this expression further simplifies to

$$\int_0^\infty \int_0^r \frac{t}{\sqrt{4\pi s^3}\sqrt{4\pi(r-s)}} \exp\left(-\frac{t^2}{4s} - \frac{||x||^2}{4r} - m^2 r\right) ds dr.$$

Doing the inner integral here, we have (via the u-substitution $u = \frac{r}{s} - 1$, hence $r - s = \frac{ru}{u+1}$ and $ds = -\frac{r}{(u+1)^2} du$)

$$\begin{aligned} \int_0^r \frac{t}{4\pi s^{3/2}\sqrt{r-s}} e^{-t^2/(4s)} ds &= \int_0^\infty \frac{t e^{-t^2(u+1)/(4r)}}{4\pi\left(\frac{r}{u+1}\right)^{3/2}\left(\frac{ru}{u+1}\right)^{1/2}} \cdot \frac{r}{(u+1)^2} du \\ &= \frac{t e^{-t^2/4r}}{4\pi r} \int_0^\infty u^{-1/2} e^{-t^2 u/(4r)} du, \end{aligned}$$

and this last thing is just a gamma integral evaluating to $\frac{\Gamma(1/2)}{\sqrt{t^2/4r}}$, so that this inner integral is exactly $\frac{e^{-t^2/(4r)}}{\sqrt{4\pi r}}$. So again substituting this back into our main expression, we have

$$G_m^0 * K_m^t = \int_0^\infty \frac{1}{(4\pi r)^{d/2}} \exp\left(-\frac{-t^2}{4r} - \frac{||x||^2}{4r} - m^2 r\right) dr,$$

and this right-hand side is exactly G_m^t as desired. □

Recall from our discussion last time that G_m^t is the amount of time d-dimensional Brownian motion spends at (t,x) when killed at a rate m, and given a point x we consider all possible y – this can be split up (based on y) as the probability that the spatial coordinate hits $x-y$ at time t and then the subsequent amount of time visited. So there's a probabilistic interpretation of this result too, but this requires the strong Markov property of Brownian motion (that is, the Markov property for stopping times rather than just fixed times).

Our next lemma will be proved more probabilistically (since the integrals are messier):

Proof of Lemma 52. We can write out the triple integral expression and then compute the $\mathbb{R}^{d-1}$ integral using the

same fact about Gaussian densities as before. We thus have

$$\begin{aligned} K_m^s * K_m^t &= \int_{\mathbb{R}^{d-1}} K_m(s,y)K_m(t,s-y) \\ &= \int_0^\infty \int_0^\infty \frac{st\exp\left(-\frac{s^2}{4u}-\frac{t^2}{4v}\right)}{u^{3/2}v^{3/2}(4\pi(u+v))^{(d-1)/2}}\exp\left(-\frac{||x||^2}{4(u+v)}-m^2(u+v)\right)dudv \\ &= \int_0^\infty \left(\int_0^r \frac{st\exp\left(-\frac{s^2}{4u}-\frac{t^2}{4(r-u)}\right)}{4\pi u^{3/2}(r-u)^{3/2}}du\right)\frac{\exp\left(-\frac{||x||^2}{4r}-m^2r\right)}{(4\pi r)^{(d-1)/2}}dr, \end{aligned}$$

where we've made the same change of variables $r = u + v$ as before. But now the inner integral is actually the convolution of two functions $p_s(u) = \frac{s}{\sqrt{4\pi u^3}}e^{-s^2/(4u)}$, $p_t(u) = \frac{t}{\sqrt{4\pi u^3}}e^{-t^2/(4u)}$. These are actually probability densities of the **Lévy distribution** of parameters s and t; that is, p_t is the probability density function of the hitting time $T_t = \inf\{s : B_s = t\}$. (This comes from the reflection principle and using the fact that the maximum of Brownian motion is distributed as $|N(0,t)|$.)

From here, the important fact is that $p_s * p_t = p_{s+t}$. We can show this for example by calculating characteristic functions, but intuitively it makes sense that the distribution of hitting time for hitting $s+t$ is the sum of hitting time for t plus the hitting time for s for any $s, t > 0$ (since we need to first hit t before $s+t$ if Brownian motion is continuous). And plugging in the expression for that new density with parameter $s+t$ yields exactly K_m^{s+t}, as desired. □

So the point is that we have indeed written our field as something at the present time plus an independent additional field, and hence $\{\phi_t\}$ is Markov. (And this really has to do with the fact that our field only interacts locally with time right before and after it because of the derivative.)

Fact 53

"Reflection positivity" can be formulated as follows: suppose we have a function of the field from time 0 onward, and now apply the same function in negative time but reversed. Then the covariance of these two functions should be nonnegative for any function. (For example if we had a real-valued process $X_1^2 + X_2$, then the covariance of that with $X_{-1}^2 + X_{-2}$ should be nonnegative.) This becomes clear from the Markov property, because if we condition on the value X_0, then this becomes a square.

So returning to our setting again, we have our Hilbert space $\mathcal{H} = L^2(\mathcal{A}, \mathcal{C}, \mu)$, where $\mathcal{A} = (\mathcal{S}(\mathbb{R}^d))^{\mathbb{R}}$, $\mathcal{C}$ is the cylinder σ-algebra, and μ is the law of the field ϕ_0 at time 0. We then define the operator $P(t)$ to act on elements of the Hilbert space $F \in \mathcal{H}$

$$P(t)F(\phi) = \mathbb{E}\left[F(\phi_t)|\phi_0 = \phi\right].$$

The question is then why this is actually a semigroup – this comes down to invariance under time-translation and also being Markov. Notice that for any times $s \le t$ and any $f, g \in \mathcal{S}(\mathbb{R}^{d-1})$, we have

$$\mathrm{Cov}(\phi_s(f), \phi_t(g)) = \int_{\mathbb{R}^{d-1}}\int_{\mathbb{R}^{d-1}} f(x)g(y)G_m(t-s, x-y)dxdy.$$

This is a function only of $(t-s)$ for any fixed f, g, and so using this identity and the fact that we have a Gaussian process, we can easily show (ϕ_s, ϕ_t) has the same distribution as (ϕ_{s+u}, ϕ_{t+u}) for any $u \in \mathbb{R}$. (Thus, our pair of $\mathcal{A}$-valued random variables has a joint law on $\mathcal{A} \times \mathcal{A}$, and if we shift the indices the joint law doesn't change – finite-dimensional distributions follow from this covariance structure for Gaussian processes.)

So now for any nonnegative s, t, we have to show that $P_{s+t}F = P_sP_tF$ (sometimes we will write subscripts P_t

instead of $P(t)$, but it means the same thing). Define the σ-algebras $\mathcal{F}_u = \sigma(\phi_u)$ and $\mathcal{F}_{\leq u} = \sigma(\{\phi_r\}_{r\leq u})$. The left-hand side can be written out as

$$\begin{aligned}\mathbb{E}\left[F(\phi_{s+t})|\phi_0\right] &= \mathbb{E}\left[F(\phi_{s+t})|\mathcal{F}_0\right] \\ &= \mathbb{E}\left[F(\phi_{s+t})|\mathcal{F}_{\leq 0}\right] \\ &= \mathbb{E}\left[\mathbb{E}\left[F(\phi_{s+t})|\mathcal{F}_{\leq s}\right]|\mathcal{F}_{\leq 0}\right]\end{aligned}$$

where we use the Markov property and then the tower property of conditional expectation. But now again by the Markov property we have $\mathbb{E}\left[F(\phi_{s+t})|\mathcal{F}_{\leq s}\right] = \mathbb{E}\left[F(\phi_{s+t})|\mathcal{F}_s\right]$; therefore **our goal is to prove that** this is equal to $P_t F(\phi_s)$, since this would show that (substituting back in)

$$\begin{aligned}P_{s+t}F &= \mathbb{E}\left[P_t F(\phi_s)|\mathcal{F}_{\leq 0}\right] \\ &= \mathbb{E}\left[P_t F(\phi_s)|\mathcal{F}_0\right] \\ &= \mathbb{E}\left[P_t F(\phi_s)|\phi_0\right] \\ &= P_s P_t F\end{aligned}$$

by the Markov property again. So to prove this last claim, take any $H \in \mathcal{H}$. We have

$$\begin{aligned}\mathbb{E}\left[P_t F(\phi_s) H(\phi_s)\right] &= \mathbb{E}\left[P_t F(\phi_0) H(\phi_0)\right] \\ &= \mathbb{E}\left[F(\phi_t) H(\phi_0)\right] \\ &= \mathbb{E}\left[F(\phi_{t+s}) H(\phi_s)\right]\end{aligned}$$

by translation-invariance in the first and third equalities, and thus by the definition of conditional expectation we indeed have equality of the blue quantities.

We'll show symmetry and strong continuity next time; the former will not be very cumbersome because G_m depends only on the squared modulus of the parameter, and the latter will require a bit more work.

15 February 10, 2025

Recall that we've been working on proving the Markov property for the process $\{\phi_t\}$. Previously, we defined

$$\psi_{s,t}(x) = \int_{\mathbb{R}^{d-1}} K_m^{t-s}(x-y)\phi_s(y)dy,$$

which is supposed to be interpreted as the "expected value of ϕ_t given the whole field at time s." It turns out that if we take any $u < s$, we have pointwise covariance (just for simplicity; we can evaluate against test functions but it'll just be messier to write out)

$$\begin{aligned}\operatorname{Cov}(\phi_u(x), \phi_t(y) - \psi_{s,t}(y)) &= G_m(t-u, x-y) - \int_{\mathbb{R}^{d-1}} K_m^{t-s}(y-z)\operatorname{Cov}(\phi_u(x), \phi_s(z))dz \\ &= G_m(t-u, x-y) - \int_{\mathbb{R}^{d-1}} K_m^{t-s}(y-z)G_m(s-u, z-x)dz\end{aligned}$$

(here in the last step notice it doesn't matter if we have $z-x$ or $x-z$ since we square anyway). And this last quantity is just $G_m(t-u, x-y) - K_m^{t-s} * G_m^s(x-y)$, but this convolution is G_m^{t-u} (by our previous lemmas) and thus this whole

covariance is zero. More generally we thus see that

$$\mathrm{Cov}(\phi_u(f), \phi_t(g) - \psi_{s,t}(g)) = 0$$

for all Schwartz functions $f, g \in \mathcal{S}(\mathbb{R}^{d-1})$. And uncorrelated implies independence for Gaussian processes (for non-Gaussians we'd have to do more), but to do things rigorously from here we have to be more precise.

Definition 54

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. We say that two sub-σ-algebras $\mathcal{G}_1, \mathcal{G}_2$ are **independent** if for all $A \in \mathcal{G}_1$ and $B \in \mathcal{G}_2$, we have $\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$. Furthermore, we say that a random variable X is independent of a σ-algebra $\mathcal{G}$ if $\sigma(X)$ and $\mathcal{G}$ are independent. Also, we say that two random variables $X, Y : \Omega \to \Omega'$ (for any measurable space $(\Omega', \mathcal{F}')$) are independent if $\sigma(X)$ and $\sigma(Y)$ are independent. (Here $\sigma(X)$ is the set of all pre-images of $\mathcal{F}'$ under X).

We can also extend this concept to collections of random variables:

Definition 55

Now let $X = \{X_i\}_{i \in I}$ be some (possibly uncountable) collection of real-valued random variables, and let $Y = \{Y_j\}_{j=J}$ be another such collection. Viewing X as an $\mathbb{R}^I$-valued random variable with the cylinder σ-algebra (and similar for Y), we say the two collections are **independent** if $\sigma(\{X_i\})$ is independent of $\sigma(\{Y_i\})$. Here, the σ-algebra generated by a collection of random variables is

$$\sigma(\{X_i\}) = \sigma\left(\left\{X_i^{-1}(A) : i \in I, A \in \mathcal{B}(\mathbb{R})\right\}\right).$$

Note that a set like "the set of ω such that $\sum |X_i|$ is finite" is not determined by a finite number of the X_is, but it is still in the cylinder σ-algebra. We can show that in fact $\sigma(\{X_i\}_{i \in I})$ is in fact exactly the same as $\sigma(X) = \{X^{-1}(A) : A \in \text{cylinder } \sigma\text{-algebra of } \mathbb{R}^I\}$.

Lemma 56

Two collections $X = \{X_i\}_{i \in I}$ and $Y = \{Y_j\}_{j \in J}$ are independent if and only if $\{X_i\}_{i \in I'}$ and $\{Y_j\}_{j \in J'}$ are independent for all **finite subsets** $I' \subseteq I$ and $J' \subseteq J$.

This is why we can get away with checking only covariances – even with infinitely many Gaussians on each side, we just need to check that random finite-dimensional vectors are independent, and for that only pairwise independence is needed.

Proof. This argument uses Dynkin's π-λ theorem twice, which will take a bit to explain. Very briefly, a collection of subsets is called a **π-system** if it is closed under intersection and a **λ-system** if it contains the full set and is closed under complements and disjoint unions. **Dynkin's π-λ theorem** then says that if $\mathcal{P}$ is a π-system, $\mathcal{L}$ is a λ-system, and $\mathcal{P} \subseteq \mathcal{L}$, then $\sigma(\mathcal{P}) \subseteq \mathcal{L}$ as well. (If $\mathcal{L}$ were fully a σ-algebra then this would be trivial, but the point is that we can relax the condition on countable unions if we have a π-system.) In many situations, **this is very useful for proving that "all Borel sets have a certain property,"** because we can just check it for intervals and then show that the sets with that property form a λ-system.

So what we're saying is that we first have independence of any $A \in \sigma(\{X_i\}_{i \in I'})$ and $B \in \sigma(\{X_j\}_{j \in J'})$ for finite I', J'; now we want to show that this also hold for B in the **full** collection. Indeed, fix A. Then $\mathcal{P} = \{B : B \in$

$\sigma(\{Y_j\}_{j\in J'})$ for some finite $J'\}$ is a π-system, and $\mathcal{L} = \{B : B \in \sigma(\{Y_j\}_{j\in J}$ and independent of $A)\}$ is a λ-system, because indeed $\mathbb{P}(\Omega \cap A) = \mathbb{P}(A) = \mathbb{P}(A)\mathbb{P}(\Omega)$, and

$$\mathbb{P}(A \cap B^c) = \mathbb{P}(A) - \mathbb{P}(A \cap B) = \mathbb{P}(A)(1 - \mathbb{P}(B)) = \mathbb{P}(A)\mathbb{P}(B^c),$$

and finally for disjoint $B_1, B_2, \cdots$ in $\mathcal{L}$, we have

$$\mathbb{P}\left(A \cap \left(\bigcup_i B_i\right)\right) = \mathbb{P}\left(\bigcup_i A \cap B_i\right) = \sum_i \mathbb{P}(A \cap B_i) = \mathbb{P}(A)\sum_i \mathbb{P}(B_i) = \mathbb{P}(A)\mathbb{P}\left(\bigcup_i B_i\right)$$

by countable additivity of probability. So by the π-λ theorem we indeed have independence for all B, not just generated by finitely many Y_is. Finally, we now turn it around and apply this argument again to show that for **any** B we can now get independence with **any** A, completing the proof. □

Remark 57. *Note that we can have collections of random variables with pairwise independence but not overall independence. But what we've just proved is that knowing independence for* **all finite** *subcollections is enough.*

Definition 58

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, and let $\mathcal{G}_0, \mathcal{G}_1, \mathcal{G}_2$ be three sub-σ-algebras. We say that $\mathcal{G}_1$ and $\mathcal{G}_2$ are **conditionally independent** of $\mathcal{G}_0$ if for all $A \in \mathcal{G}_1$, $B \in \mathcal{G}_2$, we have

$$\mathbb{P}(A \cap B|\mathcal{G}_0) = \mathbb{P}(A|\mathcal{G}_0)\mathbb{P}(B|\mathcal{G}_0) \text{ almost surely.}$$

(Here remember that these probabilities are random variables, not numbers.) We then define conditional independence of random variables in a similar way.

Definition 59

For a collection of random variables $\{X_t\}_{t\in\mathbb{R}}$, define the σ-algebras

$$\mathcal{F}_{\leq t} = \sigma\left(\{X_s\}_{s\leq t}\right), \quad \mathcal{F}_t = \sigma(X_t), \quad \mathcal{F}_{\geq t} = \sigma\left(\{X_s\}_{s\geq t}\right).$$

We say that $\{X_t\}$ is a **Markov process** if for all $t \in \mathbb{R}$, $\mathcal{F}_{\leq t}$ and $\mathcal{F}_{\geq t}$ are conditionally independent given $\mathcal{F}_t$. (In words, given the present, the past and future are conditionally independent.)

Lemma 60

Suppose we have three random variables X, Y, and Z, and suppose Z can be expressed as $F(X, T)$, where F is a measurable function and T is independent of (X, Y). Then Z and Y are conditionally independent given X.

This is what we're using to prove the Markovianity of the free field; the point is that for any fixed s, the collection $T = \{\phi_t - \psi_{s,t}\}_{t\geq s}$ is independent of $Y = \{\phi_u\}_{u\leq s}$ (because we've checked today that the necessary covariances are zero). So setting $X = \phi_s$, we indeed have T independent of (X, Y), and remembering that $\{\psi_{s,t}\}_{t\geq s}$ is a function of X only, we see that $Z = \{\phi_t\}_{t\geq s}$ is indeed a function of X and T and satisfies all of the conditions. (So we've managed to express the field at and after time s in terms of just ϕ_s and some other function T independent of anything at or before s.)

Proof of Lemma 60. We wish to show for events A, B that

$$\mathbb{P}(Z \in A, Y \in B|X) = \mathbb{P}(Z \in A|X)\mathbb{P}(Y \in B|X).$$

Unpacking this, we thus want to show that for any event C,

$$\mathbb{E}\left[\mathbb{P}(Z \in A|X)\mathbb{P}(Y \in B|X)1\{X \in C\}\right] = \mathbb{P}(Z \in A, Y \in B, X \in C),$$

so that we have a valid candidate for the conditional expectation random variable $\mathbb{P}(Z \in A, Y \in B|X)$. But writing out the probability as an integral, we have

$$\mathbb{P}(Z \in A, Y \in B, X \in C) = \int_{B\times C} \int 1\{t : F(x,t) \in A\} d\mu_T(t) \mu_{(X,Y)}(x,y).$$

On the other hand, we can prove that we actually have $\boxed{\mathbb{P}(Z \in A|X = x) = \int 1\{t : (x,t) \in A\} d\mu_T(t)}$ because the right-hand side satisfies the conditional expectation conditions for being the left-hand side. Calling this $G(x)$, we thus have

$$\begin{aligned}\mathbb{E}\left[\mathbb{P}(Z \in A|X)\mathbb{P}(Y \in B|X)1\{X \in C\}\right] &= \mathbb{E}\left[G(X)1\{X \in C, Y \in B\}\right] \\ &= \mathbb{E}\left[G(X)1\{X \in C\}\mathbb{P}(Y \in B|X)\right] \\ &= \mathbb{E}\left[\mathbb{P}(Z \in A|X)1\{X \in C\}\mathbb{P}(Y \in B|X)\right]\end{aligned}$$

where we plug in that boxed relation in the last step. And this is exactly $\mathbb{P}(Z \in A, Y \in B, X \in C)$, as desired. □

16 February 12, 2025

We've now constructed a **stationary reversible Markov process** $\{\phi_t\}_{t\in\mathbb{R}}$ on the space $(\mathcal{A}, \mathcal{C})$ for $\mathcal{A} = \mathbb{R}^{\mathcal{S}(\mathbb{R}^{d-1})}$ and $\mathcal{C}$ its cylinder σ-algebra (obtained as time-slices of the free field ϕ) – showing stationarity and reversibility are clear from looking at the finite-dimensional distributions. Our Hilbert space is then $\mathcal{H} = L^2(\mathcal{A}, \mathcal{C}, \mu)$, where μ is the law of ϕ_0, and now we want to define the semigroup $P(t) : \mathcal{H} \to \mathcal{H}$ via

$$P(t)F(\phi) = \mathbb{E}\left[F(\phi_t)|\phi_0 = \phi\right].$$

This is in fact a semigroup because of those bolded properties we've proved above, and furthermore each $P(t)$ is symmetric and $P(0)$ is the identity. So now it remains to show that $(P(t))_{t\geq 0}$ is strongly continuous.

Remark 61. *Suppose we're considering the simple harmonic oscillator on the real line. Then we can consider solutions of the Schrodinger equation starting at any state; however, if we consider the Hamiltonian for that system, it has a specific set of eigenfunctions and we get an e^{-x^2} measure. If we consider L^2 of that Gaussian measure, then we get linear combinations of Hermite polynomials, and we get evolution for those functions. So the choice of the Hilbert space is important, and we're choosing a specific one here which gives us the whole construction (if we had a function not in L^2 of the Gaussian measure, we wouldn't be able to get that). And typically in constructions like this we usually pick a ground state; in our construction here we're basically choosing μ instead and the ground state will end up being the function that is identically* 1.

Lemma 62

The map $P(t)$ defined above is a contraction; that is, $||P(t)|| \leq 1$.

Proof. This follows from Jensen's inequality for conditional expectation and then the tower property for conditional expectation:

$$\begin{aligned} ||P(t)F||^2_{L^2(\mu)} &= \mathbb{E}\left[|\mathbb{E}[F(\phi(t)|\phi_0]|^2\right] \\ &\leq \mathbb{E}\left[\mathbb{E}\left[|F(\phi(t)|^2\middle|\phi_0\right]\right] \\ &= \mathbb{E}\left[|F(\phi_t)|^2\right] \\ &= ||F||^2_{L^2(\mu)}. \end{aligned}$$

□

Recall that once we have this uniform bound on the norm on a compact interval (in fact for all t in our case), it suffices to prove strong continuity not for all F, but just for a set whose span is dense in $\mathcal{H}$. So what we'll prove is that for any $f \in \mathcal{S}(\mathbb{R}^{d-1})$, the map

$$F_f : \phi \mapsto e^{i\phi(f)}$$

will satisfy this property, and finite linear combinations of such maps are dense in $\mathcal{H}$ under the $L^2(\mu)$-topology. To show this, we'll do some machinery that may be useful for other problems we work on in the future as well.

Theorem 63

Let $(\Omega, \mathcal{F}, \mu)$ be a probability space, and let $\mathcal{A}$ be an algebra of sets that generates $\mathcal{F}$. Then given any $\varepsilon > 0$ and any $A \in \mathcal{F}$, there is some $B \in \mathcal{A}$ such that the symmetric difference satisfies $\mu(A \Delta B) < \varepsilon$.

For example, let Ω be $[0, 1)$, $\mathcal{F}$ be the Borel σ-algebra of that set, and μ Lebesgue measure, and take $\mathcal{A}$ to be the set of finite disjoint unions of half-open intervals of the form $[a, b)$ in $[0, 1)$. We can check that the empty set is in this set and that it is closed under complements, and finite unions, which is the definition of being an algebra. We can indeed check that $\sigma(\mathcal{A}) = \mathcal{B}([0, 1))$, and so the algebra $\mathcal{A}$ is easy to describe but the σ-algebra on the right-hand side is not.

Proof. Let $\mathcal{G} = \{A \in \mathcal{F} : A$ has the above property for all $\varepsilon > 0\}$ (that is, $\mathcal{G}$ contains the sets which can be arbitrarily well-approximated by things in $\mathcal{F}$). Clearly $\mathcal{G}$ contains $\mathcal{A}$, and it suffices to prove that it is a σ-algebra.

- $\mathcal{G}$ contains the emptyset because $\varnothing \in \mathcal{A}$.
- For any $A \in \mathcal{G}$ and any $\varepsilon > 0$, we can find some $B \in \mathcal{A}$ such that $\mu(A \Delta B) < \varepsilon$. But then because $A \Delta B$ is the same set as $A^c \Delta B^c$, we have $\mu(A^c \Delta B^c) < \varepsilon$, so we can also arbitrarily approximate A^c and thus $A^c \in \mathcal{G}$ as well.
- Finally, suppose we have $A_1, A_2, \cdots, \in \mathcal{G}$ and some $\varepsilon > 0$. Then we can find $B_1, B_2, \cdots \in \mathcal{A}$ such that

$\mu(A_n \Delta B_n) < \frac{\varepsilon}{2^n}$ for all n. The infinite union $B = \bigcup_{n=1}^{\infty} B_n$ may not be in $\mathcal{A}$, but note that

$$\begin{aligned}
\mu(A\Delta B) &= \mu\left(\{x : x \in A, x \notin B \text{ or } x \notin A, x \in B\}\right) \\
&\leq \mu\left(\{x : x \in A_n \text{ for some n}, x \notin B_n \text{ for all } n\}\right) + \mu\left(\{x : x \in B_n \text{ for some n}, x \notin A_n \text{ for all } n\}\right) \\
&\leq \sum_{n=1}^{\infty} \mu\left(\{x : x \in A_n, x \notin B_n\}\right) + \mu\left(\{x : x \in B_n, x \notin A_n\}\right) \\
&= \sum_{n=1}^{\infty} \mu(A_n \Delta B_n) \\
&< \varepsilon.
\end{aligned}$$

The point now is that even if B is not in $\mathcal{A}$, **because we have a finite measure**, the measure of the "tail" $\mu\left(B \setminus \bigcup_{n=1}^{N} B_n\right)$ approaches zero as $N \to \infty$. Thus there is some N with $\mu\left(A\Delta \bigcup_{n=1}^{n} B_n\right) < 2\varepsilon$, so we get arbitrary well-approximation of countable unions, as desired.

Since we have verified all conditions, we see that $\mathcal{G}$ must be all of $\sigma(\mathcal{A}) = \mathcal{F}$, as desired. □

Theorem 64

Let $(\Omega, \mathcal{F}, \mu)$ be a probability space, and let $f \in L^p(\Omega, \mathcal{F}, \mu)$ for some $p \in [1, \infty)$. Let $\mathcal{A}$ be an algebra that generates $\mathcal{F}$. Then there exists a sequence of functions $\{f_n\}$, such that each f_n is $\mathcal{A}$-measurable and $f_n \to f$ in L^p.

This is just a generalization of the previous result, and we'll use this by letting $\mathcal{F}$ be our cylinder σ-algebra and $\mathcal{A}$ the algebra of functions only dependent on finitely many coordinates. So any function in L^2 of such a space can be well-approximated by these finitely-many-coordinate functions.

To prove this, we need the following important lemma:

Lemma 65

Let $(\Omega, \mathcal{F})$ be a measurable space, and let $f : \Omega \to [0, \infty)$ be measurable functions. Then there exists a sequence of nonnegative simple functions such that $f_n \uparrow f$ pointwise.

Proof. This can be done via a direct construction: for each n, define $A_j = \{x \in \Omega : \frac{j}{2^n} \leq f(x) < \frac{j+1}{2^n}\}$ for all $0 \leq j \leq n2^n - 1$ and $A_{n2^n} = \{x \in \Omega : f(x) \geq n\}$ (that is, separate out the function into $\frac{1}{2^n}$-levels of values), and define $f_n(x) = \frac{j}{2^n}$ if $x \in A_j$. (So we truncate to $[0, n]$ and then round down to the nearest $\frac{1}{2^n}$. Notice that these sets A_j are measurable by assumption of measurability of f.) By taking powers of 2 in the denominator, we ensure that the left endpoints of our intervals always increase, and indeed the value increases pointwise to f (even if $f(x) = \infty$). □

Proof of Theorem 64. Write our function f as $f^+ - f^-$, where $f^+(x) = \max(f(x), 0)$ and $f^-(x) = \max(-f(x), 0)$; we can check that both f^+ and f^- are measurable and nonnegative. Let f_n, g_n be sequences of nonnegative simple functions that increase to f^+, f^- respectively. By the dominated convergence theorem, $|f^+ - f_n| \leq |f^+| \leq |f|$ and thus $f_n \to f^+$ in L^p. Similarly we have $g_n \to f^-$, so $f_n + g_n \to f \in L^p$. Thus if we want to prove that our collection of $\mathcal{A}$-measurable functions is dense in L^p, it suffices to show that they can approximate any simple function; in fact this means we just need to arbitrarily well-approximate any indicator function by $\mathcal{A}$-measurable functions. But we can do that by picking indicator functions of $\mathcal{A}$-measurable sets via Theorem 63. □

So now we can reduce our original problem with the quantum field to something a little simpler. Recall that $\mathcal{A} = \mathbb{R}^{\mathcal{S}(\mathbb{R}^{d-1})}$ is our space (this is not the same $\mathcal{A}$ as the algebra), $\mathcal{C}$ is our cylinder σ-algebra, and μ is the law of ϕ_0. We'll define the algebra

$$\mathcal{D} = \{C \in \mathcal{C} : C \text{ is a } \textbf{cylinder set}\},$$

where being a cylinder set means we only depend on finitely many of the coordinates; more formally it is of the form

$$C = \{\phi : (\phi(f_1), \cdots, \phi(f_n)) \in B\}$$

for some $n, f_1, \cdots, f_n$, and some Borel set $B \in B(\mathbb{R}^n)$. We can check that $\mathcal{D}$ is an algebra and that $\mathcal{C}$ is the σ-algebra that it generates, so that for any $F \in L^2(\mu)$ and any $\varepsilon > 0$, there is some $\mathcal{D}$-measurable G with $||F - G||_{L^2(\mu)} < \varepsilon$.

What we want, though, is slightly different; we want G to be a function of some finite collection of coordinates. That is, if $\mathcal{D}_{f_1,\cdots,f_n}$ is the set of all sets C of the form above with fixed $n, f_1, \cdots, f_n$ but B varying (which is a σ-algebra), we want that G is $\mathcal{D}_{f_1,\cdots,f_n}$-measurable for **some** $f_1, \cdots, f_n$, so that we can approximate by only caring about finitely many of the functions. We'll see how to bridge that next time!

17 February 14, 2025

We'll start by resolving the point from last lecture: recall that we have already proved the L^p approximation result of Theorem 64. In particular, with a small modification of the proof, we can show that the sequence $\{f_n\}$ can in fact be chosen to be made of **simple** $\mathcal{A}$-measurable functions.

Corollary 66

Let $\{(\Omega_i, \mathcal{F}_i)\}_{i \in I}$ be a collection of measurable spaces, and let $\Omega = \prod_{i \in I} \Omega_i$ be the product space and $\mathcal{F}$ be the cylinder σ-algebra. Let μ be a probability measure on $(\Omega, \mathcal{F})$, and let $f \in L^p(\mu)$ for some $p \in [1, \infty)$. Then there is a sequence $\{f_n\}$ of functions with $f_n \to f$ in L^p, and such that each f_n depends only on finitely many coordinates.

Proof. Let $\mathcal{A}$ be the set of all elements of $\mathcal{F}$ which are determined by finitely many coordinates. Then $\mathcal{A}$ is an algebra which generates the full σ-algebra (that is, $\sigma(\mathcal{A}) = \mathcal{F}$), because $\mathcal{F}$ is by definition the smallest σ-algebra containing $\mathcal{A}$. So there exist $\mathcal{A}$-measurable simple functions $f_n \to f$ in L^p, but any $\mathcal{A}$-measurable simple function can only depend on finitely many coordinates because it is a finite linear combination of indicators, which indeed each only depend on finitely many coordinates. Thus we satisfy the needed statement. □

Going back to the notation from the end of last lecture now, $\mathcal{A}$ now denotes the space $\mathbb{R}^{\mathcal{S}(\mathbb{R}^{d-1})}$ and $\mathcal{C}$ the cylinder σ-algebra. We had the Hilbert space $\mathcal{H} = L^2(\mathcal{A}, \mathcal{C}, \mu)$ (where μ is the law of ϕ_0), and what we wanted to show is that we have strong continuity of the Markov semigroup P_t. We previously proved that $||P(t)||_2 \leq 1$ for all t, so it's sufficient to show that $t \mapsto P(t)F$ is continuous for a dense subset of Fs in $\mathcal{H}$. By our corollary, it suffices to consider F depending only on finitely many coordinates.

Theorem 67

For any multivariate Gaussian distribution ν on $\mathbb{R}^n$ (of any mean and any variance), the functions $\{x \mapsto e^{it \cdot x} : t \in \mathbb{R}^n\}$ span a dense subspace of $L^2(\nu)$ (where here we mean complex L^2 rather than real L^2).

We may immediately think of Fourier transforms here, but notice that the same fact cannot be said for Lebesgue measure because $e^{it \cdot x}$ is not actually in $L^2(\text{Leb})$. So any function of finitely many jointly Gaussian variables can be

arbitrarily well-approximated by linear combinations of such exponentials; thus in our setting we have that functions of the form

$$F(\phi) = \exp\left(i(t_1\phi(f_1) + \cdots + t_n\phi(f_n))\right)$$

(for any finite $n \geq 1$, any Schwartz functions $f_1, \cdots, f_n$, and any real numbers $t_1, \cdots, t_n$) span a dense subspace of the functions depending only on **the specific functions** $f_1, \cdots, f_n$, and thus it suffices just to prove strong continuity on these Fs. But $t_1\phi(f_1) + \cdots + t_n\phi(f_n) = \phi\left(\sum_{i=1}^n t_i f_i\right)$ by linearity of the free field, and so actually it suffices to prove $t \mapsto P(t)F$ is continuous for

$$F(\phi) = e^{i\phi(f)}.$$

for any Schwartz f. To do this, we must show that for any $s \to t$, we have $P_sF \to P_tF$ in $L^2(\mu)$, meaning that

$$\mathbb{E}\left[|P_sF(\phi_0) - P_tF(\phi_0)|^2\right] \to 0.$$

But we know that

$$\begin{aligned}
\mathbb{E}\left[|P_sF(\phi_0) - P_tF(\phi_0)|^2\right] &= \mathbb{E}\left[|\mathbb{E}\left[F(\phi_t)|\phi_0\right] - \mathbb{E}\left[F(\phi_s)|\phi_0\right]|^2\right] \\
&= \mathbb{E}\left[|\mathbb{E}\left[F(\phi_t) - F(\phi_s)|\phi_0\right]|^2\right] \\
&\leq \mathbb{E}\left[\left|\mathbb{E}\left[|F(\phi_t) - F(\phi_s)|^2|\phi_0\right]\right|\right] \\
&= \mathbb{E}\left[|F(\phi_t) - F(\phi_s)|^2\right]
\end{aligned}$$

by Jensen's inequality in the third line and then the tower property. Now plugging in the expression for F, this simplifies to

$$\begin{aligned}
\mathbb{E}\left[\left|e^{i\phi_t(f)} - e^{i\phi_s(f)}\right|^2\right] &= \mathbb{E}\left[\left|e^{i(\phi_t(f) - \phi_s(f))} - 1\right|^2\right] \\
&= 2 - \mathbb{E}\left[e^{i(\phi_t(f) - \phi_s(f))}\right] - \mathbb{E}\left[e^{-i(\phi_t(f) - \phi_s(f))}\right].
\end{aligned}$$

but now because $\phi_t(f)$ and $\phi_s(f)$ are Gaussian random variables, an exact calculation (of the characteristic function of an $N(0, \sigma^2)$-distributed random variable using contour integrals) yields that this is equal to

$$2 - 2\exp\left(-\frac{1}{2}\operatorname{Var}(\phi_t(f) - \phi_s(f))\right),$$

so all we need to do is show that a certain variance $\operatorname{Var}(\phi_t(f) - \phi_s(f))$ is tending to zero as $s \to t$. Because this random variable has mean zero, we have

$$\operatorname{Var}(\phi_t(f) - \phi_s(f)) = \mathbb{E}\left[(\phi_t(f) - \phi_s(f))^2\right] = \mathbb{E}\left[\phi_t(f)^2\right] + \mathbb{E}\left[\phi_s(f)^2\right] - 2\mathbb{E}\left[\phi_t(f)\phi_s(f)\right].$$

And we know exactly how to compute each of those terms: this all simplifies to

$$2\iint f(x)f(y)G_m(0, x-y) - 2\iint f(x)f(y)G_m(t-s, x-y),$$

recalling that $\phi_t(f)$ has the same distribution as $\phi_0(f)$ because our process is stationary. And our functions are nice enough that we can apply the dominated convergence theorem to show that this indeed goes to zero as $t - s \to 0$, as desired.

So this completes the construction of the Hamiltonian – there is nothing left to prove other than our one approximation theorem.

Proof of Theorem 67. Any Gaussian random vector can be written as an affine transformation of the standard mul-

tivariate Gaussian. Thus it's sufficient to show that these maps $x \mapsto e^{it\cdot x}$ span a dense subspace of $L^2(\gamma)$ for γ the standard n-dimensional Gaussian. Furthermore, for any $f \in L^2(\gamma)$, define $f_N(x) = f(x)1\{|x| \leq N\}$. Since we have $f_N \to f$ in $L^2(\gamma)$ by the dominated convergence theorem, it suffices to consider f vanishing outside a bounded region. Take any such f (in $L^2(\gamma)$ and vanishing outside $[-N, N]^n$).

Lemma 68

For all $p \in [1, \infty)$, the Schwartz functions $\mathcal{S}(\mathbb{R}^n)$ are dense in $L^p(\mathbb{R}^n)$.

We'll take this lemma as given for now. Notice that the Gaussian density is bounded both above and below by positive numbers in a bounded region, so $f \in L^2(\gamma)$ also means $f \in L^2(\mathbb{R}^n)$, and thus there exist Schwartz functions f_k such that $f_k \to f$ in $L^2(\mathbb{R}^n)$. Again since γ has bounded density, this also means $f_k \to f$ in $L^2(\gamma)$; this it's actually sufficient to prove the theorem only for $f \in \mathcal{S}(\mathbb{R}^n)$. And now this is easy:

Lemma 69

For any $f \in \mathcal{S}(\mathbb{R}^n)$, we also have $\hat{f} \in \mathcal{S}(\mathbb{R}^n)$, and the inversion formula

$$f(x) = \frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} e^{it\cdot x} \hat{f}(t) dt$$

indeed holds.

Now $\hat{f}(t)$ is very well-behaved, and thus we can approximate this integral on the right-hand side with a finite sum of $e^{it\cdot x}$ functions to arbitrary accuracy in $L^2(\gamma)$, as desired. □

The proofs of those remaining lemmas will come later, but next on the agenda after that is to figure out "what this is good for" – it turns out that fairly generally, we can do calculations for the quantum field $U(t)$ by doing calculations for the Euclidean field $P(t)$ and doing analytic continuation. And with this, we can derive a formula for the Feynman propagator (which is used to compute basically everything and do perturbation theory). It will turn out that special relativity will "emerge naturally" from all of this!

Remark 70. *Historically, Dirac in the 1930s was trying to reconcile special relativity with quantum mechanics, but some postulate was always being violated in finite-particle systems. Instead, the idea in the 1940s was to go to these quantum fields (we'll discuss later on how we go from fields back to particles, but the key is that number of particles doesn't need to the be finite or deterministic). The first success (in the physics sense) was quantum electrodynamics (electromagnetism plus fermions), and the free field is a toy model for that. But the quantum free field didn't really appear anywhere earlier – the Euclidean free field is much simpler especially because the Green's function is actually a function (unlike the quantum one). The rigorous construction came later in the 1960s, and it is in fact important that we can do this analytic continuation carefully – we can't just put it in place of t for something like $G_m(t, x)$.*

18 February 19, 2025

We've constructed the Hilbert space and unitary group (via Stone's theorem) for the free quantum field now, and our goal is to do some calculations to check if this all matches the non-rigorous calculations done in physics. So we'll do an example, and our free field will not just be a toy example; our next goal is to quantize the classic electromagnetic field, and the construction will come from four copies of the free field.

To do the computations, we need to establish a theorem first:

Theorem 71 (Analytic continuation of semigroups)
Let $(P(t))_{t\geq 0}$ be a strongly continuous semigroup of symmetric operators on a Hilbert space $\mathcal{H}$. Let H be its generator (which we know is self-adjoint on a dense domain), and define $U(t) = e^{-itH}$ by Stone's theorem. Then for all $x, y \in \mathcal{H}$, the map

$$t \mapsto \langle P(t)x, y\rangle$$

has a (necessarily unique) analytic continuation to the open right half-plane $\Omega = \{z = x + iy : x > 0\}$ which extends continuously to $\overline{\Omega} = \{z + x + iy : x \geq 0\}$, and its value at it is exactly $\langle U(t)x, y\rangle$.

In particular, by the Riesz representation theorem, this means that we can recover $U(t)x$ for any $x \in \mathcal{H}$, since $x \mapsto \langle U(t)x, y\rangle$ is a continuous linear map for any fixed y and thus this map must be $\langle z, y\rangle$ for some z. So the point is that if we can compute those inner products and do the analytic continuation, then we can also compute with the operator $U(t)$.

The fact that we're assuming "strongly continuous semigroup of symmetric operators" is where all the analyticity comes from – these are the only assumptions we need.

Start of proof. Let $\mathcal{D}$ be the domain of the Hamiltonian of H. We'll prove this result via a sequence of lemmas:

Lemma 72
For all $t \geq 0$, $P(t)$ maps $\mathcal{D}$ into $\mathcal{D}$, and for all $t \in \mathbb{R}$, $U(t)$ also maps $\mathcal{D}$ into $\mathcal{D}$.

Proof of lemma. For any $x \in \mathcal{D}$ and $t \geq 0$, let $y = P(t)x$. Then for all $s > 0$, we have (we want to prove that this has a limit as $s \downarrow 0$)

$$\begin{aligned}\frac{P(s)y - y}{-s} &= \frac{P(s)P(t)x - P(t)x}{-s} \\ &= \frac{P(t)(P(s)x - x)}{-s}\end{aligned}$$

because $P(t)$ and $P(s)$ commute. By the Hellinger-Toeplitz theorem, $P(t)$ is continuous (because any everywhere-defined symmetric operator is continuous) and thus as $s \downarrow 0$ we have

$$\lim_{s\downarrow 0} \frac{P(t)(P(s)x - x)}{-s} = P(t) \lim_{s\downarrow 0} \frac{P(s)x - x}{-s} = P(t)Hx,$$

so the limit indeed exists and thus $P(t)x$ is also in the domain $\mathcal{D}$. The same exact proof works for $U(t)$, except that we don't even need to appeal to Hellinger-Toeplitz because we know that we have a unitary (hence continuous) operator from the start. □

Lemma 73
For all $s \geq 0$ and all $t \in \mathbb{R}$, the operators $P(s), U(t)$, and H commute on $\mathcal{D}$.

Proof of lemma. We see from the proof of the previous lemma that $P(s)$ and $U(t)$ commute with H on $\mathcal{D}$; that is, $P(s)Hx = HP(s)x$ and $U(t)Hx = HU(t)x$. Thus all that remains is to prove that $P(s)$ and $U(t)$ commute; for this we'll need to appeal to the construction of $U(t)$.

Recall that we defined $U(t)$ by considering the sequence of operators $T_n = (I + in^{-1}H)^{-1}$ (which are all well-defined on the full Hilbert space $\mathcal{H}$); moreover, T_n is a bounded linear operator from $\mathcal{H}$ to $\mathcal{D}$. We then defined $U_n(t) = e^{nt(T_n - I)} = e^{-nt}\sum_{k=0}^{\infty} \frac{(nt)^k}{k!} T_n^k$; we then found that $U_n(t)x \to U(t)x$ pointwise as $n \to \infty$.

We claim it suffices to show that $P(s)$ and T_n commute on $\mathcal{D}$. Indeed, for any $x \in \mathcal{D}$ we then have $P(s)T_n^k x = T_n P(s) T_n^{k-1} x$ (since T_n maps into the domain and hence all powers do); repeatedly moving T_ns past $P(s)$ means that $P(s)T_n^k x = T_n^k P(s)x$ for all k, so by continuity of $P(s)$ (needed because $U_n(t)$ is an infinite sum of powers of T_n) we also have $P(s)U_n(t)x = U_n(t)P(s)x$. And then again by continuity we can take $n \to \infty$, meaning that $P(s)U(t)x = U(t)P(s)x$.

So now to complete the proof of the lemma, take any $x \in \mathcal{D}$ and let $y = T_n x = (I + in^{-1}H)^{-1}x$. We have $x = (I + in^{-1}H)y$, so

$$\begin{aligned} P(s)x &= P(s)(I + in^{-1}H)y \\ &= (I + in^{-1}H)P(s)y \\ &= (I + in^{-1}H)P(s)T_n x \end{aligned}$$

because $P(s)$ commutes with H and y is in the domain. Now applying T_n to both sides cancels out the $(I + in^{-1}H)$ on the right side, which means that $T_n P(s)x = P(s)T_n x$, as desired. □

So since $P(s)$ and $U(t)$ and H all commute on the domain $\mathcal{D}$, we are ready to prove the theorem. For now, we consider any $x \in \mathcal{D}$ and $y \in \mathcal{H}$ (and then we'll prove it for all $x \in \mathcal{H}$ afterward). Let $u = s + it \in \overline{\Omega}$ (meaning that $s \geq 0$ and $t \in \mathbb{R}$); and define the operator

$$P(u) = P(s)U(t).$$

(Intuitively the idea is that e^{-uH} is defined as $e^{-sH}e^{-itH}$, since we've already worked hard to define P and U.) Defining

$$f(u) = \langle P(u)x, y\rangle,$$

we thus have $f(s) = \langle P(s)x, y\rangle$ for all $s \geq 0$ and $f(it) = \langle U(t)x, y\rangle$ for all $t \in \mathbb{R}$. Our goal is therefore to prove that f is analytic on Ω and continuous on $\overline{\Omega}$; this would give us the analytic continuation.

We first need to prove differentiability of f at every point $u = s + it \in \Omega$, and this will use strong continuity. For fixed x and any $\varepsilon \geq 0$, define the functions

$$R(\varepsilon) = ||P(\varepsilon)x - x + \varepsilon Hx||, \quad S(\varepsilon) = ||U(\varepsilon)x - x + i\varepsilon Hx||,$$

$$A(\varepsilon) = ||P(\varepsilon)Hx - Hx||, \quad B(\varepsilon) = ||U(\varepsilon)Hx - Hx||.$$

All of these are well-defined because we're assuming x is in the domain. Also because x is in the domain, $R(\varepsilon), S(\varepsilon)$ are $o(\varepsilon)$ as $\varepsilon \to 0$ (because we know that if we divide these by ε they go to zero), and also by strong continuity $A(\varepsilon), B(\varepsilon)$ are $o(1)$ as $\varepsilon \to 0$.

Lemma 74

There is some constant C (depending on s) such that for all $a, b \in [0, s+1]$, we have

$$||P(b)x - P(a)x + (b-a)HP(a)x|| \leq CR(|a-b|) + C|a-b|A(|a-b|).$$

for the functions R, A above; in particular this quantity on the left-hand side must be $o(|a-b|)$.

Proof of lemma. First suppose $b \geq a$. Then

$$||P(b)x - P(a)x + (b-a)HP(a)x|| = ||P(a)\left(P(b-a)x - x + (b-a)Hx\right)|\,,|$$

and because we've proven that (by the uniform boundedness principle) the norms of $P(a)$ are bounded on the compact interval $[0, s+1]$ by some C, this norm is at most $C||P(a)\left(P(b-a)x - x + (b-a)Hx\right)|| = CR(|b-a|)$, as desired.

Now if $b < a$, we need to do a bit of extra work: the triangle inequality yields

$$||P(b)x - P(a)x + (b-a)HP(a)x|| \leq ||P(a)x - P(b)x - (a-b)HP(b)x|| + |a-b| \cdot ||P(b)Hx - P(a)Hx||.$$

The first term is bounded, as before, by $CR(|b-a|)$ (because we've now swapped the roles of a and b). And for the second term, we write

$$||P(a)Hx - P(b)Hx|| = ||P(b)(P(a-b)Hx - Hx)|| \leq CA(|a-b|),$$

so plugging this back in gives us the desired result. □

A similar proof also gives us the following lemma (but now remembering that U is uniformly bounded by 1):

Lemma 75

For all $a, b \in \mathbb{R}$, we have

$$||U(b)x - U(a)x|| + i(b-a)HU(a)x|| \leq S(|a-b|) + |a-b|B(|a-b|).$$

To prove our result now, consider any $u' = s' + it' \in \Omega$ with $s' \in (0, s+1]$. We now try to differentiate at s: we have

$$\begin{aligned} P(u')x - P(u)x &= P(s')U(t')x - P(s)U(t)x \\ &= P(s')(U(t') - U(t))x + U(t)(P(s') - P(s))x, \end{aligned}$$

where it's okay to subtract $P(s')U(t)x$ and then add $U(t)P(s')x$ because $U(t), P(s')$ map into the domain and commute on the domain. We can then compute

$$\begin{aligned} |||P(u')x - P(u)x + (u'-u)HP(u)x|| &= ||P(s')(U(t') - U(t))x + U(t)(P(s') - P(s))x + (((s'-s) + i(t'-t))HP(u)x)|| \\ &= ||P(s')(U(t') - U(t))x + U(t)(P(s') - P(s))x + (s'-s)HU(t)P(s)x \\ &\quad + i(t'-t)HP(s)U(t)x|| \end{aligned}$$

(notice that we've written $P(u)$ as $P(s)U(t)$ in one of these expressions and $U(t)P(s)$ in the other). We can now decompose the blue $P(s)$ as $P(s) - P(s') + P(s')$, and then furthermore we can combine the red parts and also the green parts so that this whole expression simplifies to

$$\begin{aligned} |||P(u')x - P(u)x + (u'-u)HP(u)x|| \leq\ & ||P(s')(U(t') - U(t) + i(t'-t)HU(t))x|| \\ & + ||U(t)(P(s') - P(s) + (s'-s)HP(s))x|| \\ & + |t'-t| \cdot ||U(t)(P(s') - P(s))Hx||. \end{aligned}$$

Next time, we'll show that all of these terms are $o(u'-u)$ so that we do indeed have differentiability, and then Cauchy-Riemann will yield analyticity for any $x \in \mathcal{D}$. Then to go from $x \in \mathcal{D}$ to $x \in \mathcal{H}$, we will use the well-known result that "if a sequence of analytic functions defined on an open domain converges pointwise to a limit, and we have uniform convergence on compact sets, then the limit is also analytic." □

19 February 21, 2025

Last time, we were proving Theorem 71 and produced a candidate for the analytic continuation, namely $P(u) = P(s)U(t)$ for any $u = s + it$ with $s \geq 0$. (What's nice is that Stone's theorem already gives us $U(t)$, so all we're proving is that the analytic continuation is actually consistent with that.) We showed that $P(s)$ and $U(t)$ all map into the domain $\mathcal{D}$ and that they commute with each other and with H. (Remember that even though H is defined on a dense subset, both the group U and the semigroup P are defined on **all** of $\mathcal{H}$.) We'll continue the rest of our proof now:

Proof of Theorem 71, continued. Recall that we're trying to prove that $f(u) = \langle P(u)x, y\rangle$ is analytic in our region Ω. For this, we just need to show that f is complex differentiable at u, and we're doing this by brute-force methods – our goal will be to show that, as expected,

$$f'(u) = -\langle HP(u)x, y\rangle.$$

Specifically, this means we must show that

$$||f(u') - f(u) + (u' - u)\langle HP(u)x, y\rangle|| = o(u' - u)$$

for complex numbers $u' \to u$. For now, we're considering only the case $x \in \mathcal{D}, y \in \mathcal{H}$. But by Cauchy-Schwarz,

$$\begin{aligned} ||f(u') - f(u) + (u' - u)\langle HP(u)x, y\rangle|| &= |\langle P(u')x - P(u)x + (u' - u)HP(u)x, y\rangle| \\ &\leq ||P(u')x - P(u)x + (u' - u)HP(u)x|| \cdot ||y||, \end{aligned}$$

so it suffices to show that the first norm in the last expression is $o(|u' - u|)$. We decomposed that norm at the end of last lecture into three terms – recall that for any $u = s + it$ and $u' = s' + it'$ with $s, s' > 0$ and $s' \leq s + 1$, we have

$$\begin{aligned} |||P(u')x - P(u)x + (u' - u)HP(u)x|| \leq\ & ||P(s')(U(t') - U(t) + i(t' - t)HU(t))x|| \\ &+ ||U(t)(P(s') - P(s) + (s' - s)HP(s))x|| \\ &+ |t' - t| \cdot ||U(t)(P(s') - P(s))Hx||. \end{aligned}$$

But now recall the definitions of $R(\varepsilon), S(\varepsilon), A(\varepsilon), B(\varepsilon)$ from last lecture and their associated bounds in Lemma 74 and Lemma 75. We can use these to bound each of the terms above: we know that $||U(t)|| = 1$ and $||P(s')|| \leq C$ for some constant C (on the interval $[0, s + 1]$), so in fact

$$|||P(u')x - P(u)x + (u' - u)HP(u)x|| \leq C\left(S(|t' - t|) + |t' - t|B(|t' - t|)\right) + CR(|s' - s|) + C|s' - s|A(|s' - s|) + |t' - t|$$

where on the last term, we are using that $P(s) - P(s') = P(s')(P(s - s') - I)$ if $s \geq s'$ and then applying the lemma for P (or doing it the other way around if $s < s'$). And now because $R(\varepsilon), O(\varepsilon)$ are $o(\varepsilon)$ and $A(\varepsilon), B(\varepsilon)$ are $o(1)$, all of these terms are either $o(|s' - s|)$ or $o(|t' - t|)$, so we've proven that differentiability does indeed hold for any $f(u)$ with x in the domain.

Remark 76. *There are more sophisticated ways of proving differentiability, such as Morera's theorem. But for this problem, we don't need such sophisticated things, since we can get the derivative directly.*

So now to complete the proof of analyticity, take a general $x \in \mathcal{H}$ and let $x_n \in \mathcal{D}$ be a sequence of points such

that $x_n \to x$. For any $u \in \Omega$, define $f_n(u) = \langle P(u)x_n, y\rangle$ and $f(u) = \langle P(u)x, y\rangle$. Then

$$\begin{aligned} |f_n(u) - f(u)| &= |\langle P(u)(x_n - x), y\rangle| \\ &= ||P(u)(x_n - x)|| \cdot ||y||, \end{aligned}$$

and now for any compact subset $K \subseteq \Omega$, we have $\sup_{u\in K} ||P(u)||$ finite, since this quantity is at most $\sup_{u\in K} ||P(\mathrm{Re}(u))||$ (and we know that $||P(t)||$ is uniformly bounded on compact subsets for real t). So in fact $f_n \to f$ uniformly on compact sets, and therefore f is also analytic (this can be proved by Morera's theorem, since indeed integrals along any triangle will be zero).

Finally, for continuous extension to the imaginary axis, take any $u = it$ for $t \in \mathbb{R}$. We know that

$$f(u) = \langle P(it)x, y\rangle = \langle U(t)x, y\rangle$$

by definition; now let $u_n = s_n + it_n$ be a sequence of points with $t_n \to t$ and s_n positive and converging to zero. We want to show that $f(u_n) \to f(u)$, but this is the same proof strategy as before: we have

$$|f(u_n) - f(u)| \leq ||(P(u_n) - P(u))x|| \cdot ||y||$$

by Cauchy-Schwarz and properties of the operator norm, and now

$$\begin{aligned} ||(P(u_n) - P(u))x|| &= ||P(s_n)U(t_n) - U(t)x|| \\ &\leq ||P(s_n)(U(t_n) - U(t))x|| + ||(P(s_n) - I)U(t)x||. \end{aligned}$$

The first term is bounded by $C||U(t_n) - U(t)x||$, and so now by strong continuity of the semigroup and group both terms indeed go to zero, as desired. □

Remark 77. *It is in fact important that P is a semigroup, not a group. In general, because we work in L^2 but then immediately get very smooth functions when applying $P(t)$, we can't "go backwards" and define $P(-t)$ except for things that are very smooth. And this also explains why we can't analytically extend our function to negative real part.*

We'll now apply this theorem to compute some "vacuum expectation values" for the systems we're considering. Let's first summarize everything up to this point and recall our setting: we wish to construct a quantum system taking values in some space $\mathcal{X}$ endowed with a σ-algebra. So if we have some trajectory $\gamma : [t_0, t_1] \to \mathcal{X}$, physics gives us some action $S(\gamma)$, which we assume has the additive property (so if γ is the concatenation of two trajectories γ_1, γ_2, then $S(\gamma) = S(\gamma_1) + S(\gamma_2)$). Recall that this additive property is what gives us Markovianity. The steps that we must take are **(1)** obtain the Euclidean action S_E by Wick rotation, **(2)** rigorously define a stationary reversible $\mathcal{X}$-valued Markov process which formally has probability density proportional to $e^{-S_E(\gamma)}$ (the definition can be rigorous, though the justification of the density must be heuristic), **(3)** set μ to be the law of X_0 and define $\mathcal{H} = L^2(\mu)$, and **(4)** define the semigroup via

$$P(t)F(x) = \mathbb{E}\left[F(X_t)|X_0 = x\right].$$

By stationarity, reversibility, and the Markov property, we know that $(P(t))_{t\geq 0}$ is a symmetric semigroup, and so **(5)** we must prove on a case-by-case basis that it is strongly continuous. If that holds, then we get a quantum system by the mechanism above.

20 February 24, 2025

We'll prove a theorem about contracting semigroups today before diving into vacuum expectations:

Definition 78

Let $\mathcal{H}$ be a Hilbert space and $(P(t))_{t\geq 0}$ a strongly continuous semigroup of symmetric operators on $\mathcal{H}$. Then the semigroup is called **contracting** if $||P(t)|| \leq 1$ for all $t \geq 0$.

All of the semigroups we have been constructing via Markov processes have this property.

Theorem 79

With the notation in the above definition, let H be the generator of $(P(t))$. Then $\langle Hx, x\rangle \geq 0$ for all $x \in \mathcal{D}$ (that is, we have a positive operator) and the spectrum of H is contained in $[0, \infty)$.

Intuitively, $P(t) = e^{-tH}$, so if there were an element of the spectrum below zero then this operator wouldn't be contracting.

Proof. Since we already know H is self-adjoint, its spectrum is contained in the real line, and we just need to show negative numbers are not in the spectrum. Recall the definition of this for densely defined operators: let $\mathcal{D}$ be the domain of H; for any $\alpha < 0$, we have to show that $H - \alpha I : \mathcal{D} \to \mathcal{H}$ is invertible (that is, it is a linear bijection) and its inverse is bounded (that is, continuous). Indeed, we can explicitly construct the inverse: set

$$R = \int_0^\infty e^{\alpha t} P(t) dt,$$

noting that this integral is well-defined by a Riemann sum approximation because $||P(t)|| \leq 1$ and α is negative. This R is a linear operator from $\mathcal{H}$ into $\mathcal{H}$, and by the triangle inequality

$$||R|| \leq \int_0^\infty e^{\alpha t} ||P(t)|| dt \leq \int_0^\infty e^{\alpha t} dt < \infty$$

so that R is bounded. We will show that this is indeed the inverse of $H - \alpha I$ (which is intuitively what should happen if we think of P as e^{-tH}). This is similar to what we've done many times: by manipulating the terms, we can show that for any $x \in \mathcal{H}$,

$$(P(s) - I)Rx = (e^{-\alpha s} - 1)\left(Rx - \int_0^s e^{\alpha t} P(t) x dt\right) - \int_0^s e^{\alpha t} P(t) x dt,$$

and now if we divide both sides by $-s$ and take the limit as $s \downarrow 0$, $e^{-\alpha s} - 1$ will become α an the last term approaches x, so in fact

$$\lim_{s \downarrow 0} \frac{(P(s) - I)Rx}{-s} = \alpha Rx + x.$$

So this shows that Rx is indeed in the domain, so that R maps from $\mathcal{H}$ into $\mathcal{D}$, and also that $HRx = \alpha Rx + x$, which rearranges to $(H - \alpha I)Rx = x$. Therefore $H - \alpha I : \mathcal{D} \to \mathcal{H}$ is surjective.

Finally, to show that it is also injective (which will show that we do indeed have an inverse), we have by Cauchy-Schwarz that for all $x \in \mathcal{H}$,

$$\langle P(t)x, x\rangle \leq ||P(t)x|| \cdot ||x|| \leq ||x||^2 = \langle x, x\rangle,$$

so that (subtracting the right expression from the left and dividing by $-t$)

$$\left\langle \frac{P(t)x - x}{-t}, x \right\rangle \geq 0.$$

Taking a limit as $t \downarrow 0$, this shows that $\langle Hx, x\rangle \geq 0$ for all $x \in \mathcal{D}$ (as desired). And furthermore if $x \in \ker(H - \alpha I)$, then $Hx = \alpha x$, but this means that $\langle \alpha x, x\rangle = \langle Hx, x\rangle \geq 0$. In other words, $\alpha ||x||^2 \geq 0$, but this can only occur if $x = 0$ because α is negative. Therefore $H - \alpha I : \mathcal{D} \to \mathcal{H}$ is a linear bijection. Together with knowing that R is a one-sided inverse, we see that it is indeed a two-sided inverse, completing the proof. □

Returning to our setting, we have a probability space $(\mathcal{A}, \mathcal{C}, \mu)$ and the associated Hilbert space $H = L^2(\mathcal{A}, \mathcal{C}, \mu)$. We then have an $\mathcal{A}$-valued (stationary, reversible) Markov process $\{X_t\}_{t\in\mathbb{R}}$, and we defined the semigroup via conditional expectations

$$P(t)F(x) = \mathbb{E}\left[F(X_t)|X_0 = x\right].$$

Jensen's inequality for conditional expectation shows that $||P(t)F||_{L^2(\mu)} \leq ||F||_{L^2(\mu)}$, so $(P(t))_{t\geq 0}$ is a contracting semigroup. This means that we know our spectrum is nonnegative, and in fact we do have a "ground state" corresponding to the minimum point 0 in the spectrum: for the function Ω which is identically one, we have $P(t)\Omega = \Omega$ for all t, so $H\Omega = 0$. (Notice that the stationary measure μ is already accounted for here.) Usually this is the unique ground state, but we won't assume that (uniqueness requires a separate proof); we'll just have this be the canonical ground state.

Fact 80

Any element of $\mathcal{H}$ is indeed a wavefunction relative to the μ-measure; the interpretation is that if we're in a state F, then for all $A \in \mathcal{C}$, the probability that our quantum system belongs to A is

$$\mathbb{P}(\phi \in A) = \frac{\int_A |F(x)|^2 d\mu(x)}{\int_{\mathcal{A}} |F(x)|^2 d\mu(x)}.$$

For example, in the simple harmonic oscillator, μ is the Gaussian measure, and our eigenfunctions are just Hermite polynomials.

Example 81

We'll now think about expectations using the path integral formulation: for this, we'll go back to the one-particle system for simplicity. Suppose we want to evaluate the expected value of a quantity like

$$F_1(\gamma(t_1))F_2(\gamma(t_2))\cdots F_k(\gamma(t_k))$$

for some ordered set of times $t_1 < \cdots < t_k$ and some functions $F_1, \cdots, F_k$, **given that** the particle starts in state ψ_0 at some time $-T < t_1$ and ends at state ψ_1 at time $T > t_k$. (It might not be entirely clear rigorously what this means, but intuitively we're given conditions on the particle at some distant past and future time.) Notice also that these are observables F which only depend on the position of the particle.

This setting has some (not mathematically precise) prescription in the path integral framework: the expected value (or matrix element) is given by

$$\langle \psi_0 | F_1(\gamma(t_1))\cdots F_k(\gamma(t_k))|\psi_1\rangle = \frac{1}{Z(2T)} \int \psi_0(\gamma(-T))\overline{\psi_1(\gamma(T))}F_1(\gamma(t_1))\cdots F_k(\gamma(t_k))e^{iS(\gamma)}\mathcal{D}\gamma$$

for some normalizing constant $Z(2T)$. But in the framework of what we've been doing, we can transfer from Lebesgue measure to $L^2(\mu)$, where $d\mu(x) = f(x)dx$ for some probability density f, and then a state ψ now becomes $\tilde{\psi} = \frac{\psi}{\sqrt{f}}$. If we take the expression above and write it in terms of our transformed states, we now have

$$\langle \tilde{\psi}_0 | F_1(\gamma(t_1)) \cdots F_k(\gamma(t_k)) | \tilde{\psi}_1 \rangle = \frac{1}{Z(2T)} \int \tilde{\psi}_0(\gamma(-T)) \overline{\tilde{\psi}_1(\gamma(T))} \sqrt{f(\gamma(-T)) f(\gamma(T))} F_1(\gamma(t_1)) \cdots F_k(\gamma(t_k)) e^{iS(\gamma)} \mathcal{D}\gamma.$$

If we simplify and let $\tilde{\psi}_0$ and $\tilde{\psi}_1$ both be the ground state Ω (that is, identically 1), we get

$$\langle \Omega | F_1(\gamma(t_1)) \cdots F_k(\gamma(t_k)) | \Omega \rangle = \frac{1}{Z(2T)} \int \sqrt{f(\gamma(-T)) f(\gamma(T))} F_1(\gamma(t_1)) \cdots F_k(\gamma(t_k)) e^{iS(\gamma)} \mathcal{D}\gamma,$$

and our goal now is to evaluate this. (In the physics phrasing, we want to evaluate elements of the S matrix, where the idea is that particles "come in to a collision initially as free particles, scatter according to interactions, and then eventually act as free particles again," and we want to understand the scattering probabilities.) The path integral formulation says that the time-evolution of our state is given by

$$\tilde{U}(t)\tilde{\psi}(x) = \frac{1}{Z(t)\sqrt{f(x)}} \int_{\substack{\gamma:[0,t]\to\mathbb{R} \\ \gamma(0)=x}} \psi(\gamma(t)) \sqrt{f(\gamma(t))} e^{iS(\gamma)} \mathcal{D}\gamma.$$

If we wanted to evaluate the expectation for just a single function (so $k = 1$), then we have

$$\begin{aligned} \langle \Omega | F_1(\gamma(t_1)) | \Omega \rangle &= \frac{1}{Z(2T)} \int \sqrt{f(\gamma(-T)) f(\gamma(T))} F_1(\gamma(t_1)) e^{iS(\gamma)} \mathcal{D}\gamma \\ &= \frac{1}{Z(2T)} \int_{\substack{\gamma_1:[-T,t_1]\to\mathbb{R} \\ \gamma_2:[t_1,T]\to\mathbb{R} \\ \gamma_1(t_1)=\gamma_2(t_1)}} \sqrt{f(\gamma_1(-T)) f(\gamma_2(T))} \frac{\sqrt{f(\gamma_1(t_1))}}{\sqrt{f(\gamma_1(t_1))}} F_1(\gamma(t_1)) e^{iS(\gamma_1)} e^{iS(\gamma_2)} \mathcal{D}\gamma_1 \mathcal{D}\gamma_2 \end{aligned}$$

where now we're assuming that the action is additive and where the red term is just multiplying by 1. We can also check that $Z(T_1 + T_2) = Z(T_1)Z(T_2)$ here, so if we now fix γ_1 and evaluate the inner integral on γ_2, we actually just have

$$\frac{1}{Z(T - t_1)\sqrt{f(\gamma_1(t_1))}} \int_{\substack{\gamma_2:[t_1,T]\to\mathbb{R} \\ \gamma_2(t_1)=\gamma_1(t_1)}} e^{iS(\gamma_2)} \sqrt{f(\gamma_2(T))} \mathcal{D}\gamma_2 = \tilde{U}(T - t_1)\Omega = \Omega$$

is identically 1, because it's the path integral formula applied to the wavefunction Ω. This in particular tells us that the normalizing constant is given by

$$Z(t) = \frac{1}{\sqrt{f(x)}} \int_{\substack{\gamma:[0,t]\to\mathbb{R} \\ \gamma(0)=x}} \sqrt{f(\gamma(t))} e^{iS(\gamma)} \mathcal{D}\gamma$$

(and in fact the right-hand side is independent of x), and therefore we can get rid of the contributions to the path integral after t_1:

$$\langle \Omega | F_1(\gamma(t_1)) | \Omega \rangle = \frac{1}{Z(t_1 + T)} \int_{\gamma_1:[-T,t_1]\in\mathbb{R}} \sqrt{f(\gamma_1(-T)) f(\gamma_1(t_1))} F_1(\gamma_1(t_1)) e^{iS(\gamma_1)} \mathcal{D}\gamma_1,$$

noting now that there are no restrictions on the starting point of γ_1. We can now calculate this, since we want to get rid of path integrals completely and just write things in terms of the unitary evolution: we may write this integration as an integral over the starting point separately, so that we have

$$\int_{-\infty}^{\infty} f(x) \frac{1}{Z(t_1 + t)\sqrt{f(x)}} \int_{\gamma:[-T,t_1\to\mathbb{R}\gamma(-T)=x} \sqrt{f(\gamma(t_1))} F_1(\gamma(t_1)) e^{iS(\gamma)} \mathcal{D}\gamma dx$$

but by the path integral formulation, the red part is just $\tilde{U}(t_1+T)F_1(x)$. Thus this whole thing evaluates to

$$\langle\Omega|F_1(\gamma(t_1))|\Omega\rangle = \int_{-\infty}^{\infty} \tilde{U}(t_1+T)F_1(x)d\mu(x),$$

and now this is actually just $\int F_1(x)d\mu(x)$, since in the space $L^2(\mu)$ we're trying to evaluate the inner product

$$\langle\Omega, \tilde{U}(t_1+T)F\rangle_{L^2(\mu)},$$

and now the adjoint of $\tilde{U}(t_1+T)$ is just $\tilde{U}(-t_1-T)$, which leaves Ω unchanged. So we don't have to do a lot of work to compute the expectations for just a single function – one way of summarizing this is that we're computing $\Omega e^{itH}Fe^{-itH}\Omega$, but this just becomes $\Omega F\Omega$.

Next time we'll see what we have to do with multiple functions and multiple times: for example with two functions, it turns out we'll have to apply $U(t_2-t_1)$ to F_2 and take expectations and so on.

21 February 26, 2025

We'll continue the discussion from last time: recall that our quantum field takes values in some space $\mathcal{X}$, and we have some additive action S on (finite or infinite) trajectories of the field. We also have some heuristic notion of "Lebesgue measure" λ on $\mathcal{X}$, allowing us to do the path integral, and our original Hilbert space is "$\mathcal{H}_0 = L^2(\mathcal{X},\lambda)$" (this doesn't quite make sense, but it's what we're trying to make sense of). We know that a state $\psi\in\mathcal{H}_0$ evolves as

$$U(t)\psi(x) = \frac{1}{Z(t)}\int_{\substack{\gamma:[0,t]\to\mathcal{X}\\ \gamma(0)=x}} \psi(\gamma(t))e^{iS(\gamma)}\mathcal{D}\gamma,$$

and (just to review) to rigorously construct this, we first obtain a Euclidean action S_E and construct a probability measure on the space of trajectories $\mathcal{X}^{\mathbb{R}}$ that formally has density proportional to $e^{-S_E(\gamma)}$ with respect to "Lebesgue measure" on $\mathcal{X}^{\mathbb{R}}$. Defining an $\mathcal{X}$-valued stochastic process $\{X_t\}_{t\in\mathbb{R}}$ with this probability distribution, and supposing that we get a stationary reversible Markov process, we then move the whole thing to a new Hilbert space $\mathcal{H}=L^2(\mu)$, where μ is the law of X_0 (or any X_t). The state $\psi\in\mathcal{H}_0$ then corresponds to the state $\tilde{\psi}=\frac{\psi}{\sqrt{f}}$ in $\mathcal{H}$, where $f=\frac{d\mu}{d\lambda}$ is the density of our stationary measure, and the evolution now takes the form

$$\tilde{U}(t)\tilde{\psi}(x) = \frac{1}{Z(t)\sqrt{f(x)}}\int_{\substack{\gamma:[0,t]\to\mathcal{X}\\ \gamma(0)=x}} \tilde{\psi}(\gamma(t))\sqrt{f(\gamma(t))}e^{iS(\gamma)}\mathcal{D}\gamma$$

where $\mathcal{D}\gamma$ is still "integration with respect to Lebesgue measure in each coordinate" (except the first which is fixed). And what we've recently proved is that we can obtain $\tilde{U}$ by defining $P(t)\psi(x)=\mathbb{E}[\psi(X_t)|X_0=x]$ and letting $\tilde{U}=e^{-itH}$ for the generator H of $(P(t))_{t\geq 0}$.

Fact 82

Here the Hilbert space, $P(t)$ and $\tilde{U}(t)$ are rigorously defined objects, so what we construct is fully rigorous – what's not rigorous is the justification that "this does the job for quantum field theory." But as long as we can reproduce the calculations, that's a justification that we're doing the right thing.

Remark 83. *Physicists work with interacting quantum systems in four dimensions, but no one has been able to rigorously construct such an object yet. (And here, it turns out that "interacting" basically just means "non-Gaussian.")*

We'll now write down a formula for the vacuum expectations and justify it. Recall that the vacuum state $\Omega\in L^2(\mu)$ is the state which is identically 1, and because $P(t)$ is a conditional expectation, $P(t)\Omega=\Omega$ for all t and thus

(differentiating) $H\Omega = 0$. (Since our semigroup is contracting, the spectrum is contained in $[0, \infty)$ and thus this attains the minimum eigenvalue.)

Remark 84. *Note that the vacuum state in our original Hilbert space is $\Omega(x) = \sqrt{f(x)}$, and only when we move to the new one do we have $\tilde{\Omega}(x) = 1$.*

This also means that Ω doesn't evolve under $\tilde{U}(t) = e^{-itH}$, but that means

$$\begin{aligned} 1 = \Omega(x) &= \tilde{U}(t)\Omega(x) \\ &= \frac{1}{Z(t)\sqrt{f(x)}} \int_{\substack{\gamma:[0,t]\to\mathcal{X} \\ \gamma(0)=x}} \Omega(\gamma(t)) e^{iS(\gamma)} \mathcal{D}\gamma \\ &= \frac{1}{Z(t)\sqrt{f(x)}} \int_{\substack{\gamma:[0,t]\to\mathcal{X} \\ \gamma(0)=x}} \sqrt{f(\gamma(t))} e^{iS(\gamma)} \mathcal{D}\gamma, \end{aligned}$$

meaning that $Z(t)$ can be expressed as a path integral

$$\boxed{Z(t) = \frac{1}{\sqrt{f(x)}} \int_{\substack{\gamma:[0,t]\to\mathcal{X} \\ \gamma(0)=x}} \sqrt{f(\gamma(t))} e^{iS(\gamma)} \mathcal{D}\gamma}$$

for any x. This tells us that

$$\begin{aligned} Z(s+t) &= \frac{1}{\sqrt{f(x)}} \int_{\substack{\gamma:[0,s+t]\to\mathcal{X} \\ \gamma(0)=x}} \sqrt{f(\gamma(s+t))} e^{iS(\gamma)} \mathcal{D}\gamma \\ &= \frac{1}{\sqrt{f(x)}} \int_{\substack{\gamma_1:[0,s]\to\mathcal{X} \\ \gamma_1(0)=x}} \int_{\substack{\gamma_2:[s,s+t]\to\mathcal{X} \\ \gamma_2(s)=\gamma_1(s)}} \sqrt{f(\gamma_2(s+t))} e^{iS(\gamma_1)+iS(\gamma_2)} \mathcal{D}\gamma_2 \mathcal{D}\gamma_1, \end{aligned}$$

but now the inner integral is exactly $Z(t)\sqrt{f(\gamma_1(s))}$, and then pulling that out of the integrand leaves just $Z(s)$, so that $Z(s+t) = Z(t)Z(s)$ as desired.

Remark 85. *In particular, this tells us that even if we go back to the Lebesgue measure space, Z would still be multiplicative.*

Furthermore if we take the boxed equation, multiply both sides by $f(x)$, and integrate over x, we have $\int f(x)\lambda(x) = 1$ and thus we find that

$$Z(t) = \int_{\gamma:[0,t]\to\mathcal{X}} \sqrt{f(\gamma(0))f(\gamma(t))} e^{iS(\gamma)} \mathcal{D}\gamma$$

(where now $\mathcal{D}\gamma$ also includes "integrating over $d\lambda(\gamma(0))$," even though it originally didn't because we were fixing the starting location). We'll use this for the next justification: our goal is now to define, for ϕ our quantum field, the quantity

$$\langle \Omega | F_1(\phi_{t_1}) F_2(\phi_{t_2}) | \Omega \rangle ,$$

where $F_1, F_2 : \mathcal{X} \to \mathbb{C}$ are functions. The heuristic definition of this is that (now using the original Hilbert space, so $\Omega(x) = \sqrt{f(x)}$ here)

$$\begin{aligned} \langle \Omega | F_1(\phi_{t_1}) F_2(\phi_{t_2}) | \Omega \rangle &= \lim_{T\to\infty} \langle \Omega | F_1(\phi_{t_1}) F_2(\phi_{t_2}) | \Omega \rangle_T \\ &= \frac{1}{Z(2T)} \int_{\gamma:[-T,T]\to\mathcal{X}} \Omega(\gamma(-T))\overline{\Omega(\gamma(T))} F_1(\gamma(t_1)) F_2(\gamma(t_2)) e^{iS(\gamma)} \mathcal{D}\gamma \end{aligned}$$

for any $-T < t_1 < t_2 < T$. (Notice that if we plug in $F_1 = F_2 = 1$ into this expression, the integral matches with the

expression for Z and thus this "expectation value" is indeed 1.) Rewriting in our new Hilbert space, we have

$$\frac{1}{Z(2T)}\int_{\gamma:[-T,T]\to\mathcal{X}}\sqrt{f(\gamma(-T))f(\gamma(T))}F_1(\gamma(t_1))F_2(\gamma(t_2))e^{iS(\gamma)}\mathcal{D}\gamma$$
$$=\frac{1}{Z(2T)}\int_{\gamma_1:[-T,t_1]\to\mathcal{X}}\int_{\substack{\gamma_2:[t_1,t_2]\to\mathcal{X}\\ \gamma_2(t_1)=\gamma_1(t_1)}}\int_{\substack{\gamma_3:[t_2,T]\to\mathcal{X}\\ \gamma_3(t_2)=\gamma_2(t_2)}}\sqrt{f(\gamma_1(-T))f(\gamma_3(T))}F_1(\gamma_1(t_1))F_2(\gamma_2(t_2))e^{iS(\gamma_1)}e^{iS(\gamma_2)}e^{iS(\gamma_3)}mcD\gamma_3\mathcal{D}\gamma_2\mathcal{D}\gamma_1$$

after breaking up the path into three parts. Integrating over γ_3, this simplifies to

$$\int_{\gamma_1:[-T,t_1]\to\mathcal{X}}\int_{\substack{\gamma_2:[t_1,t_2]\to\mathcal{X}\\ \gamma_2(t_1)=\gamma_1(t_1)}}F_1(\gamma_1(t_1))F_2(\gamma_2(t_2))e^{iS(\gamma_1)}e^{iS(\gamma_2)}\sqrt{f(\gamma_2(t_2))}Z(T-t_2)\mathcal{D}\gamma_2\mathcal{D}\gamma_1$$

by the boxed formula above for the normalizing constant. Next, integrating over γ_2, we're left with

$$\frac{Z(T-t_2)}{Z(2T)}\int_{\gamma_1:[-T,t_1]\to\mathcal{X}}\sqrt{f(\gamma_1(-T))}F(\gamma_1(t_1))e^{iS(\gamma_1)}Z(t_2-t_1)\tilde{U}(t_2-t_1)F_2(\gamma_1(t_1))\sqrt{f(\gamma_1(f_1))}\mathcal{D}\gamma_1$$

by the definition of our evolution $\tilde{U}$. And finally we can separate out the integral over the initial point for γ_1: defining $G_1(x)=F_1(x)\tilde{U}(t_2-t_1)F_2(x)$, we're left with

$$\frac{Z(T-t_2)Z(t_2-t_1)}{Z(2T)}\int_{\mathcal{X}}\sqrt{f(x)}\int_{\substack{\gamma_1:[-T,t_1]\to\mathcal{X}\\ \gamma_1(0)=x}}\sqrt{f(\gamma_1(t_1))}G_1(\gamma_1(t_1))e^{iS(\gamma_1)}\mathcal{D}\gamma_1 d\lambda(x).$$

If we now divide and multiply by $\sqrt{f(x)}$ again, the expression we end up with is

$$\frac{Z(T-T_2)Z(t_2-t_1)Z(t_1-(-T))}{Z(2T)}\int_{\mathcal{X}}f(x)G_0(x)d\lambda(x),$$

where $G_0=\tilde{U}(t_1-(-T))G_1$ and that first fraction exactly cancels out by multiplicativity of Z. Since $f(x)d\lambda(x)$ is just $d\mu$, we see that this whole thing just evaluates to $\int_{\mathcal{X}}G_0(x)d\mu(x)$, and now here G_0 is a well-defined quantity (start with F_2, apply some operator to it, multiply F_1, apply another operator). But we can do a further simplification, since

$$\begin{aligned}\int_{\mathcal{X}}G_0(x)d\mu(x)&=\langle G_0,\tilde{\Omega}\rangle_{L^2(\mu)}\\&=\left\langle\tilde{U}(t_1-(-T))G_1,\Omega\right\rangle\\&=\left\langle G_1,\tilde{U}(-t_1+(-T))\tilde{\Omega}\right\rangle\\&=\left\langle G_1,\tilde{\Omega}\right\rangle\\&=\int_{\mathcal{X}}G_1(x)d\mu(x).\end{aligned}$$

So the point is that we just integrate the function $G_1=F_1\tilde{U}(t_2-t_1)F_2$, which does not depend on T as long as t_1,t_2 are both contained in $[-T,T]$! Therefore it's possible to make the following rigorous definition:

Proposition 86

The vacuum expectation value is

$$\left\langle\Omega\middle|F_1(\phi_{t_1})F_2(\phi_{t_2})\middle|\Omega\right\rangle=\int_{\mathcal{X}}G_1(x)d\mu(x),$$

where $G_1=F_1\tilde{U}(t_2-t_1)F_2$ is the product of two functions in $L^2(\mu)$, meaning that $G_1\in L^1$ and the integral makes sense.

We have to be a bit more careful if we have more than two functions in our expectation – the iterative things we do have to all be in L^2. But there's a more general theorem we can get there too. And next time, our goal will be to compute this using analytic continuation.

Remark 87. *When we compute expectations "given the far past and far future," we often want things to not depend on those conditions. But this usually requires Markov chains to be ergodic, and that would require its own proof. What's nice is that the way we do things here actually does not depend on our sufficiently large T.*

22 February 28, 2025

We'll derive the Feynman propagator for the free field today – this is the foundation of all calculations we do in more complicated settings, since most calculations are perturbations around free fields (or Grassmann fields for fermions, though it's not known how to make that setting precise at all).

Theorem 88

Let φ be a (quantum, **note for clarity going forward the notation change** compared with the Euclidean version ϕ) free field on $\mathbb{R}^4$ with mass m. For $x, y \in \mathbb{R}^4$, we have

$$\langle \Omega | \mathcal{T}(\varphi(x)\varphi(y)) | \Omega \rangle = -i\Delta_F(x-y),$$

where $\mathcal{T}$ is the **time-ordering operator**

$$\mathcal{T}(\varphi(x)\varphi(y)) = \begin{cases} \varphi(x)\varphi(y) & x_0 \le y_0, \\ \varphi(y)\varphi(x) & y_0 \le x_0, \end{cases}$$

and Δ_F is the **Feynman propagator** defined by (now we switch notation to writing points in $\mathbb{R}^4$ as (t, x) for $t = x_0$ and $x = (x_1, x_2, x_3)$)

$$\Delta_F(t,x) = \frac{i}{(2\pi)^3} \int_{\mathbb{R}^3} \frac{e^{-i|t|\omega_p + ip\cdot x}}{2\omega_p} dp, \quad \omega_p = \sqrt{||p||^2 + m^2}.$$

Note that the expression for Δ_F is not integrable, so instead it should be interpreted as a distribution integrated against functions. And in particular that means the left-hand side (the two-point correlation) is also a distribution. Gaussian fields are determined by two-point correlations, but here φ is the quantum free field rather than the Euclidean one, so we're not actually working with Gaussian objects here. Instead, we have to do things via analytic continuation, and this Feynman propagator will allow us to define expectations of various functions instead of "defining φ directly."

Remark 89. *Special relativity comes out of this – Δ_F is Lorentz-invariant, which is clear if we rewrite it in the more familiar physics form*

$$\Delta_F = \lim_{\varepsilon \downarrow 0} \frac{1}{(2\pi)^3} \int_{\mathbb{R}^4} \frac{e^{i(x,p)}}{p^2 + m^2 - i\varepsilon} dp$$

where x now represents all four coordinates and $(x,p) = -x_0p_0 + x_1p_1 = x_2p_2 + x_3p_3$ and $p^2 = -p_0^2 + p_1^2 + p_2^2 + p_3^2$ now come from the Minkowski metric, meaning that we indeed get an expression that is manifestly Lorentz-invariant.

To prove this formula, we'll take two points $(s,x), (t,y)$ with $s \le t$ and $x, y \in \mathbb{R}^3$. We want to show that the expectation

$$\langle \Omega | \varphi(s,x)\varphi(t,y) | \Omega \rangle$$

evaluates to something, and first we need to figure out what it means. If we take two Schwartz functions $f, g \in \mathcal{S}(\mathbb{R}^3)$, we can define the functions **on our Euclidean space** $L^2(\mathcal{A}, \mathcal{C}, \mu)$

$$F_1(\phi) = \phi(g), \quad F_2(\phi) = \phi(f).$$

So if we multiply by $g(x)$ and $f(y)$ and integrate, we find that formally

$$\int_{\mathbb{R}^3}\int_{\mathbb{R}^3} f(y)g(x)\,\langle\Omega|\varphi(s,x)\varphi(t,y)|\Omega\rangle\,dxdy = \langle\Omega|F_1(\varphi_s)F_2(\phi_t)|\Omega\rangle\,,$$

where $\phi_s(x) = \phi(s,x)$ (so that we're applying F_1 and F_2 to timeslices). If we had the Euclidean free field ϕ in place of φ here, this would make sense rigorously, but now we want to do it for the quantum free field. By our previous discussion, our definition of this quantity is actually

$$\langle\Omega|F_1(\varphi_s)F_2(\varphi_t)|\Omega\rangle = \int_{\mathcal{A}} F_1(\phi)U_{t-s}F_2(\phi)d\mu(\phi),$$

where μ is the law of ϕ_0 and $(U_t)_{t\in\mathbb{R}}$ is the evolution group for the quantum free field.

Remark 90. *If we did this whole calculation with the ordinary inner product instead of the "physics" one $\langle\Omega|\cdot|\Omega\rangle$, then we would end up with the same expression but with P_{t-s} instead of U_{t-s}. So this is just like going from the heat kernel to the Schrodinger kernel, and indeed the Schrodinger kernel is much less friendly to work with than the heat kernel.*

We'll now match up the two integral expressions to derive our expression for the correlation function. Taking $s = 0$ for simplicity, we then want to compute $\langle U_t F_2, \overline{F_1}\rangle_{L^2(\mu)}$, and we know the way to evaluate this: we define

$$h(t) = \langle P_t F_2, \overline{F}_1\rangle_{L^2(\mu)}$$

for all real $t \geq 0$, and then analytically continue h to the right half-plane and extend by continuity to the imaginary axis, evaluating $h(it)$ for $t \in \mathbb{R}$. But we know that for real t,

$$h(t) = \mathbb{E}[F_1(\phi_0)P_tF_2(\phi_0)] = \mathbb{E}[F_1(\phi_0)F_2(\phi_t)] = \mathbb{E}[\phi_0(g)\phi_t(f)],$$

so finally we have something we can actually evaluate. Recall that this last expression is exactly

$$h(t) = \int_{\mathbb{R}^3}\int_{\mathbb{R}^3} f(y)g(x)G_m(t,x-y)dxdy, \quad \text{where} \quad G_m(t,x) = \int_0^\infty \frac{1}{(4\pi t)^2}\exp\left(-\frac{t^2}{4\tau} - \frac{||x-y||^2}{4\tau} - m^2\tau\right)d\tau.$$

If we substitute this in, we now want to do an analytic continuation in t. And since t only appears inside the exponential, it would be bad to just replace t with it – that would give us a divergent integral. So we have to do something more clever:

Lemma 91

For $u, v \in \mathcal{S}(\mathbb{R}^4)$, we have the alternative expression in terms of Fourier transforms

$$\int_{\mathbb{R}^4}\int_{\mathbb{R}^4} u(x)v(y)G_m(x-y)dxdy = \frac{1}{(2\pi)^4}\int_{\mathbb{R}^4}\frac{\hat{u}(p)\overline{\hat{v}(p)}}{||p||^2+m^2}dp.$$

The point is that analytically continuing in terms of Fourier transforms will be made easier once the ω_p factor comes out, and it's related to this thing called the "mass shell."

Proof of lemma. We have to do some smoothing: we consider the quantity

$$\iint u(x)v(y)e^{-\varepsilon||x||^2}G_{m+\varepsilon}(x-y)dxdy$$

as $\varepsilon \to 0$ (which converges to the expression we want by dominated convergence theorem). After doing the Fourier transforms, we get (we need to regularize m by ε as well so that Fubini's theorem applies)

$$\frac{1}{(2\pi)^8}\iiiint \hat{u}(p)\overline{\hat{v}(q)}e^{-ip\cdot x+iq\cdot y-\varepsilon||x||^2}\frac{1}{(4\pi t)^2}\exp\left(-\frac{||x-y||^2}{4t}-(m+\varepsilon)^2 t\right)dtdpdqdxdy,$$

and now integrating over y is just a Gaussian integral, yielding

$$\frac{1}{(2\pi)^8}\iiiint \hat{u}(p)\overline{\hat{v}(q)}e^{-i(p-q)\cdot x-\varepsilon||x||^2-((m+\varepsilon)^2+||q||^2)t}dtdpdqdx.$$

Next t only appears in one place, so we can integrate out over that to get

$$\frac{1}{(2\pi)^8}\iiint \frac{\hat{u}(p)\overline{\hat{v}(q)}}{||q||^2+(m+\varepsilon)^2}e^{-i(p-q)\cdot x-\varepsilon||x||^2}dqdpdx.$$

Finally the integral over x is another Gaussian integral, so we end up with

$$=\frac{(\pi/\varepsilon)^2}{(2\pi)^8}\iint \frac{\hat{u}(p)\overline{\hat{v}(q)}}{||q||^2+(m+\varepsilon)^2}\exp\left(-\frac{1}{4\varepsilon}||p-q||^2\right)dpdq,$$

and then after a change of variables and taking $\varepsilon \to 0$ we get the desired result. □

So this is the rigorous version of saying that "$\delta_0(x) = \frac{1}{(2\pi)^d}\int_{\mathbb{R}^d} e^{-ip\cdot x}dp$;" we just need to regularize with the Gaussian term.

We next want to do the same thing but "with a delta mass on the time coordinate:"

Lemma 92

We have for any real-valued Schwartz functions $f, g \in \mathcal{S}(\mathbb{R}^3)$ that

$$\int_{\mathbb{R}^3}\int_{\mathbb{R}^3} f(x)g(y)G_m(t,x-y)dxdy = \frac{1}{(2\pi)^3}\int_{\mathbb{R}^3}\frac{e^{-t\omega_p}}{2\omega_p}\hat{f}(p)\overline{\hat{g}(p)}dp,$$

with $\omega_p = \sqrt{||p||^2+m^2}$.

Start of proof. Define

$$f_\varepsilon(s,x) = f(x)\frac{e^{-\frac{1}{2\varepsilon}(s-t)^2}}{\sqrt{2\pi\varepsilon}}, \quad g_\varepsilon(s,x) = g(x)\cdot\frac{e^{-\frac{1}{2\varepsilon}s^2}}{\sqrt{2\pi\varepsilon}}.,$$

We'll basically use this in our previous lemma and then send $\varepsilon \to 0$; we'll end up getting a similar to integral to what we want but with an extra variable to integrate out. □

23 March 3, 2025

We'll continue the proof of Lemma 92 today:

Proof of Lemma 92, continued. Recall that the functions $f_\varepsilon(s,x) = f(x)\frac{e^{-\frac{1}{2\varepsilon}(s-t)^2}}{\sqrt{2\pi\varepsilon}}$ and $g_\varepsilon(s,x) = g(x)\frac{e^{-\frac{1}{2\varepsilon}s^2}}{\sqrt{2\pi\varepsilon}}$ from the end of last lecture are meant to be f and g in the space coordinates, multiplied by approximations of a delta mass in the time coordinate at t and 0, respectively. We can easily check that $f_\varepsilon, g_\varepsilon$ are in $\mathcal{S}(\mathbb{R}^4)$. On the one hand, we have by the dominated convergence theorem that as $\varepsilon \to 0$,

$$\iiiint f_\varepsilon(s,x)g_\varepsilon(r,y)G_m(s-r,x-y)\,dxdydrds \to \iint f(x)g(y)G_m(t,x-y)dxdy,$$

but on the other hand we have the Fourier expression of f_ε

$$\hat{f}_\varepsilon(q,p) = \iint e^{isq} e^{ix\cdot p} f(x) \frac{e^{-\frac{1}{2\varepsilon}(s-t)^2}}{\sqrt{2\pi\varepsilon}} ds dx,$$

which simplifies to $\hat{f}(p)e^{itq-\frac{1}{2}\varepsilon q^2}$ because the s-integral is just a Gaussian integral. Similarly, we find that $\hat{g}_\varepsilon(q,p) = \hat{g}(p)e^{-\frac{1}{2}\varepsilon q^2}$. Thus applying Lemma 91, we have

$$\iiiint f_\varepsilon(s,x) g_\varepsilon(r,y) G_m(s-r, x-y)\, dxdydrds = \frac{1}{(2\pi)^4} \iint \frac{\hat{f}_\varepsilon(q,p)\overline{\hat{q}_\varepsilon(q,p)}}{q^2 + ||p||^2 + m^2} dqdp$$

(where q is real and p is in $\mathbb{R}^3$), and plugging in our expressions for the Fourier transforms yields

$$\frac{1}{(2\pi)^4} \iint \frac{e^{itq-\varepsilon q^2} \hat{f}(p)\overline{\hat{g}(p)}}{q^2 + ||p||^2 + m^2} dqdp,$$

and now again by the dominated convergence theorem this has a limit as $\varepsilon \to 0$, which is

$$\begin{aligned}\int_{\mathbb{R}^3}\int_{\mathbb{R}^3} f(x)g(y)G_m(t,x-y)dxdy &= \lim_{\varepsilon\to 0} \iiiint f_\varepsilon(s,x) g_\varepsilon(r,y) G_m(s-r,x-y)\, dxdydrds \\ &= \frac{1}{(2\pi)^4} \iint \frac{e^{itq}\hat{f}(p)\overline{\hat{g}(p)}}{q^2 + ||p||^2 + m^2} dqdp.\end{aligned}$$

Note that we can't yet simplify by replacing t with it and get something sensible for our analytic continuation yet – we have to do one more step. Integrating out q, we use the fact that $\int_{-\infty}^{\infty} \frac{e^{isx}}{x^2+a^2} dx = \frac{\pi}{a} e^{-a|s|}$ (obtained by contour integration and the residue theorem); we may recognize this as the characteristic function for the Cauchy distribution. Therefore (we have $a^2 = ||p||^2 + m^2$)

$$\int_{\mathbb{R}^3}\int_{\mathbb{R}^3} f(x)g(y)G_m(t,x-y)dxdy = \frac{1}{(2\pi)^3} \int_{\mathbb{R}^3} \frac{e^{-t\sqrt{||p||^2+m^2}}}{2\sqrt{||p||^2+m^2}} \hat{f}(p)\overline{\hat{g}(p)} dp,$$

as desired. (And in words, this is the Fourier transform of the heat kernel for the Euclidean free field.) □

If we define $F_1(\phi) = \phi(g)$ and $F_2(\phi) = \phi(f)$, this quantity we are evaluating is $\langle P_t F_1, F_2\rangle_{L^2(\mu)}$, which is the same as $\mathbb{E}[\phi_0(g)\phi_t(f)]$. So we are ready to do the analytic continuation because we're now in a better situation: the expression

$$\psi(z) = \frac{1}{(2\pi)^3} \int_{\mathbb{R}^3} \frac{e^{-z\sqrt{||p||^2+m^2}}}{2\sqrt{||p||^2+m^2}} \hat{f}(p)\overline{\hat{g}(p)}$$

is well-defined and analytic in the open right half-plane (because we get an exponential suppression with a positive real part, and f and g are Schwartz). It further extends continuously to the imaginary axis by continuation, so that we have

$$\psi(it) = \frac{1}{(2\pi)^3} \int_{\mathbb{R}^3} \frac{e^{-it\sqrt{||p||^2+m^2}}}{2\sqrt{||p||^2+m^2}} \hat{f}(p)\overline{\hat{g}(p)},$$

and so this must be exactly the quantity $\langle U(t)F_2, F_1\rangle$ we were after. Expanding out the expressions for the Fourier coefficients, we thus get

$$\langle U(t)F_2, F_1\rangle = \frac{1}{(2\pi)^3} \iiint g(x)f(y) \frac{e^{-it\omega_p + ip\cdot(x-y)}}{2\omega_p} dxdydp,$$

and then we can extract the Feynman propagator by comparing this with the **heuristic** expression

$$\langle U(t)F_2, F_1\rangle_{L^2(\mu)} = \langle \Omega | F_1(\phi_0)F_2(\phi_t)|\Omega\rangle = \iint g(x)f(y)\, \langle\Omega|\varphi_0(x)\varphi_t(y)|\Omega\rangle\, dxdy;$$

we see that the expression we get by integrating out p is what we should get for $\langle\Omega|\varphi_0(x)\varphi_t(y)|\Omega\rangle$.

Remark 93. *Recall that the relativistic energy of a particle is $\sqrt{||p||^2c^2+m^2c^4}$, which is exactly our ω_p in units where $c=1$. And when we're moving much slower than the speed of light, meaning $p=mv$ has $||v|| \ll c$, this is approximately $mc^2+\frac{1}{2}m||v||^2$, which is the rest energy plus the kinetic energy. But our relativistic form is Lorentz-invariant. In contrast, in the Schrodinger equation we have a term $\frac{\partial^2}{\partial x^2}$ which corresponds to p^2 in Fourier space, but we should also be including the mc^2 part to make it relativistic – the Fourier transform should really evolve as*

$$\widehat{i\frac{\partial\psi}{\partial t}} = \sqrt{||p||^2c^2+m^2c^4}\hat{\psi},$$

but this "instantaneously gives information" and thus breaks the rule that nothing should travel faster than the speed of light. (For example if $\hat{\psi}$ starts off as a delta function, then instantaneously we get something that's nonzero everywhere.) So as we've been saying, we'll need to work with fields instead, and the state of the field can also describe the state of a system of particles. ($L^2(\mu)$ can be identified with a bosonic Fock space.)

Fact 94

If we want to use a negative time instead, we have to alter our time-ordering: $\mathcal{T}(\varphi_0(g)\varphi_t(f))$ for negative t will have the two terms be swapped, and so that's why we get a $|t|$ in the actual expression. (These two-point correlation functions are only defined when time-ordered.)

We'll now do the full construction of the free electromagnetic field; this is basically using four independent free fields and thus should be relatively simple from all of the work we've been doing. (But there are some complications that come here as well.) Recall that we have an electric field $\vec{E}=(E_1,E_2,E_3):\mathbb{R}^4\to\mathbb{R}^3$ and magnetic field $\vec{B}=(B_1,B_2,B_3):\mathbb{R}^4\to\mathbb{R}^3$ at each point in spacetime, obtained as follows: we have a gauge field $A=(A_0,A_1,A_2,A_3):\mathbb{R}^4\to(i\mathbb{R})^4$ (imaginary axis because this is a Lie-algebra valued form – this is a different convention from what's used in physics) which relates to $\vec{E}$ and $\vec{B}$ through the formulas

$$E_j = i\frac{\partial A_j}{\partial t} - i\frac{\partial A_0}{\partial x_j}$$

for $j=1,2,3$, and

$$B_1 = i\frac{\partial A_2}{\partial x_3} - i\frac{\partial A_3}{\partial x_2}, \quad B_2 = i\frac{\partial A_1}{\partial x_3} - i\frac{\partial A_3}{\partial x_1}, \quad B_3 = i\frac{\partial A_2}{\partial x_1} - i\frac{\partial A_1}{\partial x_2}.$$

The pair $(\vec{E},\vec{B})$ is the physical object, while A is the mathematical construct; the map $A\mapsto(\vec{E},\vec{B})$ is many-to-one, and two gauge fields A, A' give rise to the same $(\vec{E},\vec{B})$ if one is a gauge transform of the other, meaning that

$$A_0' = A_0 + i\frac{\partial\psi}{\partial t}, \quad A_j' = A_j + i\frac{\partial\psi}{\partial x_j}$$

for some $\psi:\mathbb{R}^4\to\mathbb{R}$. Next lecture, we'll quantize A (so we can lift from Maxwell's equations to quantum field theory).

24 March 5, 2025

Today, we'll start with our object of interest $A=(A_0,A_1,A_2,A_3):\mathbb{R}^4\to(i\mathbb{R})^4$ and quantize it. We observed previously that the action $S(A)$ is invariant under gauge transformations $A\mapsto A^\psi$; we'll use this invariance to transform the problem of quantization to something more manageable (via fixing the Feynman gauge).

Definition 95

A $U(1)$ gauge field A satisfies the **Lorenz gauge condition** if we have everywhere that

$$-\frac{\partial A_0}{\partial t}+\sum_{j=1}^{3}\frac{\partial A_j}{\partial x_j}=0.$$

The idea of Feynman gauge fixing is that each gauge equivalence class contains a **unique element** satisfying the Lorenz gauge condition, and any field in the class is mapped to this element by a unique gauge transformation. We'll need to do some work to make this precise, since we've said nothing about the gauge field or transforms other that they are smooth, and eventually we'll be working with distributions (not even functions). So this is all for the sake of justifying the heuristics; we'll proceed with heuristics to arrive at a model, and we'll define that model rigorously.

With this point of view, we can think of a gauge field A as a pair $(\tilde{A},\psi)$, where $\tilde{A}$ is a gauge field satisfying the Lorenz condition and ψ is a gauge transformation. The action of $(\tilde{A},\psi)$ then only depends on $\tilde{A}$ (since $S(A)=S(\tilde{A})$), so from the path integral formulation we see that replacing the above action with $\tilde{S}(\tilde{A},\psi)=S(\tilde{A})+F(\psi)$ for **any** function F has no effect on the theory. (This is like a probability model where we have a pair of random variables but the density is uniform in the latter; basically we're keeping the independence of the two coordinates and not changing the distribution of the first.)

The justification we need to make is that the expectation of any gauge-invariant observable is the same when we do this (since all physically interesting observables $\mathcal{O}$ should satisfy $\mathcal{O}(A^\psi)=\mathcal{O}(A)$). More precisely, we're saying that vacuum expectations $\left\langle\Omega\middle|\mathcal{O}_1(A(t_1)\cdots\mathcal{O}_n(A_{t_n)}\middle|\Omega\right\rangle$ should be the same under S and $\tilde{S}$ if $\mathcal{O}_1,\cdots,\mathcal{O}_n$ are gauge-invariant. We can see that this is true, because when we do the Wick rotation, the Euclidean action $S_E(\tilde{A})$ gets replaced by $S_E(\tilde{A})+F_E(\psi)$. The probability measure was originally proportional to $\exp(-S_E(\tilde{A}))$, and under the new action it is instead proportional to $\exp(-S_E(\tilde{A})-F_E(\psi))$; thus the distribution of $\tilde{A}$ is not changing and thus when we do the steps for computing the vacuum expectation value on the Euclidean side, the $\mathcal{O}_i$s are not affected by the choice of F, and thus the analytic continuation is not affected either.

We just need to be careful to choose our F so that the probability measure does actually make sense (for example if it's a quadratic function of ψ with certain conditions on the coefficients), and we'll make a specific choice and develop the theory this way. But first we'll prove the uniqueness:

Proposition 96

Let $\mathcal{A}_S$ be the space of all Schwartzian $U(1)$ connections of $\mathbb{R}^4$ (meaning each component is a Schwartz function), let $\mathcal{A}_S^L$ be all such connections satisfying the Lorenz gauge condition, and let $\mathcal{G}_S$ be the set of all Schwartzian gauge transformations. Define the map $\Phi:\mathcal{A}_S^L\times\mathcal{G}_S\to\mathcal{A}_S$ by $\Phi(A,\psi)=A^\psi$. Then Φ is bilinear and injective, and its image is dense in $\mathcal{A}_S$ in the L^2 topology.

Proof. Linearity of the map Φ is clear from the definition of the gauge transformation. Thus to show injectivity, we just need to show that the kernel is trivial. Suppose $(A,\psi)\in\ker(\Phi)$; then we must have

$$\frac{\partial\psi}{\partial t}=iA_0,\quad \frac{\partial\psi}{\partial x_j}=iA_j$$

for $j=1,2,3$. Then the operator

$$\Box\psi=\frac{\partial^2\psi}{\partial t}-\sum_{j=1}^{3}\frac{\partial^2\psi}{\partial x_j^2}=i\frac{\partial A_0}{\partial t}-i\sum_{j=1}^{3}\frac{\partial A_j}{\partial x_j}=0$$

by the Lorenz gauge transformation. Applying the Fourier transform (valid because we have Schwartz functions), we have

$$\widehat{\Box\psi}(p) = (p_0^2 - p_1^2 - p_2^2 - p_3^2)\hat{\psi}(p) = 0,$$

which implies that $\hat{\psi} = 0$ almost everywhere on $\mathbb{R}^4$ (except the 3-dimensional mainfold where $p_0^2 = p_1^2 + p_2^2 + p_3^2$); since Schwartz functions are continuous this means $\hat{\psi}$ (which is also Schwartz) must be identically zero, so $\psi = 0$ and therefore $A = 0$, as desired.

Finally, for density, take any Schwartzian connection $A \in \mathcal{A}_S$ and define

$$\phi = \frac{\partial A_0}{\partial t} - \sum_{j=1}^{3} \frac{\partial A_j}{\partial x_j}.$$

Then ϕ is Schwartz, and define

$$g(p) = \frac{i\hat{\phi}(p)}{p_0^2 - p_1^2 - p_2^2 - p_3^2}$$

whenever the denominator is nonzero. If we then define the inverse Fourier transform

$$\psi(x) = \frac{1}{(2\pi)^4} \int_{\mathbb{R}^4} g(p) e^{-ip\cdot x} dp,$$

the problem is that in general this integral may not converge unless we assume g extends continuously to a Schwartz function also on the manifold $\{p : p_0^2 - p_1^2 - p_2^2 - p_3^2 = 0\}$ (because the integral may not be absolutely convergent). But if it does, then $\hat{\psi} = g$, and we claim $\tilde{A} = A^{\psi}$ satisfies the Lorenz gauge condition. Indeed, from the definition we have

$$\frac{\partial \tilde{A}_0}{\partial t} - \sum_{j=1}^{3} \frac{\partial \tilde{A}_j}{\partial x_j} = \phi + i\Box\psi,$$

and it just remains to show that the right-hand side is zero. But

$$\widehat{\Box\psi}(p) = (p_0^2 - p_1^2 - p_2^2 - p_3^2)\hat{\psi}(p) = (p_0^2 - p_1^2 - p_2^2 - p_3^2)g(p) = i\hat{\phi}(p),$$

so $\hat{\phi}(p) + i\widehat{\Box\psi}(p) = 0$ and thus $\phi + i\Box\psi = 0$ as well and thus the right-hand side is indeed zero. And the class of all A for which this g is nicely behaved enough is indeed dense (multiplying by a sufficient factor when we're close to that manifold). □

Remark 97. *We're basically trying to solve the wave equation for Schwartz functions here – what we really want is to define*

$$\psi(x) = \frac{1}{(2\pi)^4} \int_{\mathbb{R}^4} e^{-ipx} \frac{i\hat{\phi}(p)}{p_0^2 - p_1^2 - p_2^2 - p_3^2} dp.$$

The question is whether this is well-defined if A is Schwartz, and also whether the end result is Schwartz.

Either way, we're ready to make our specific choice of F now: our new action $\tilde{S}(A) = S(\tilde{A}) + F(\psi)$ takes the **Feynman gauge**

$$F(\psi) = \frac{1}{2} \int_{\mathbb{R}} \int_{\mathbb{R}^3} (\Box\psi(t,x))^2 \, dx dt.$$

(We'll see why this particular form, up to the constant, is nice.) We then have

$$\begin{aligned} S(\tilde{A}) + F(\psi) &= S(A) + \frac{1}{2}\int_{\mathbb{R}}\int_{\mathbb{R}^3} (\Box\psi(t,x))^2\, dxdt \\ &= S(A) - \frac{1}{2}\int_{\mathbb{R}}\int_{\mathbb{R}^3} (i\Box\psi)^2\, dxdt \\ &= S(A) - \frac{1}{2}\int_{\mathbb{R}}\int_{\mathbb{R}^3} \left(\frac{\partial \tilde{A}_0}{\partial t} - \sum_{j=1}^{3}\frac{\partial \tilde{A}_j}{\partial x_j} + i\Box\psi\right)^2 dxdt \\ &= S(A) - \frac{1}{2}\int_{\mathbb{R}}\int_{\mathbb{R}^3} \left(\frac{\partial A_0}{\partial t} - \sum_{j=1}^{3}\frac{\partial A_j}{\partial x_j}\right)^2 dxdt \end{aligned}$$

(second-to-last line because the thing we add is just zero). So the point is we're back to something that just depends on A but without changing the expectation values, and next time we'll see why this is easier to work with!

25 March 7, 2025

Last time, we replaced the original action $S(A)$ by a modified action $S(A) - \frac{1}{2}\int_{\mathbb{R}}\int_{\mathbb{R}^3}\left(\frac{\partial A_0}{\partial t} - \sum_{j=1}^{3}\frac{\partial A_j}{\partial x_j}\right)^2 dxdt$, and we will now try to construct the system with this action. Heuristically "by integration by parts" (denoting t by x_0) we have

$$\iint \frac{\partial A_j}{\partial x_k}\frac{\partial A_k}{\partial x_j}dxdt = \iint \frac{\partial A_j}{\partial x_j}\frac{\partial A_k}{\partial x_k}dxdt,$$

so if we expand out the square in our action and define the vector $\eta = (-1,1,1,1)$, we have

$$\begin{aligned} &-\frac{1}{2}\int_{\mathbb{R}^4}\sum_{0\le j<k\le 3}\eta_j\eta_k\left(\frac{\partial A_k}{\partial x_j} - \frac{\partial A_j}{\partial x_k}\right)^2 dx \\ &= -\frac{1}{4}\int_{\mathbb{R}^4}\sum_{0\le j,k\le 3}\eta_j\eta_k\left(\frac{\partial A_k}{\partial x_j} - \frac{\partial A_j}{\partial x_k}\right)^2 dx \\ &= -\frac{1}{4}\int_{\mathbb{R}^4}\left(\sum_{j,k}\eta_j\eta_k\left(\frac{\partial A_k}{\partial x_j}\right)^2 + \sum_{j,k}\eta_j\eta_k\left(\frac{\partial A_j}{\partial x_k}\right)^2 - 2\sum_{j,k}\eta_j\eta_k\left(\frac{\partial A_k}{\partial x_j}\frac{\partial A_j}{\partial x_k}\right)\right) dx \\ &= -\frac{1}{2}\int_{\mathbb{R}^4}\left(\sum_{j,k}\eta_j\eta_k\left(\frac{\partial A_k}{\partial x_j}\right)^2 - \left(\sum_{j}\eta_j\frac{\partial A_j}{\partial x_j}\right)^2\right) dx, \end{aligned}$$

where the first step is rewriting our integral in the new convention, the second step is because the $j = k$ terms contribute nothing and so we can symmetrize over $j < k$ and $j > k$, the third step is expanding out squares, and the fourth step is reindexing the first two sums and applying the integration by parts formula on the last sum. But now the gauge-fixing term cancels out (this is why the factor of $\frac{1}{2}$ was important), and so we end up with our expression for the action being

$$-\frac{1}{2}\int_{\mathbb{R}^4}\sum_{j,k=0}^{3}\eta_j\eta_k\left(\frac{\partial A_k}{\partial x_j}\right)^2 dx = -\frac{1}{2}\int_{\mathbb{R}}\int_{\mathbb{R}^3}\left((\partial_t A_0)^2 - ||\nabla A_0||^2\right)dxdt + \frac{1}{2}\sum_{k=1}^{3}\int_{\mathbb{R}}\int_{\mathbb{R}^3}\left((\partial_t A_k)^2 - ||\nabla A_k||^2\right)dxdt$$

where we're switching back to our original notation and writing two different terms depending on whether $k = 0$ or $k > 1$. And now this is looking a bit more manageable – it's closer to four copies of the free field – so now we can do a Wick rotation. We should replace dt with $-idt$, and ∂_t should then be replaced by $i\partial_t$; making these replacements

in both times, we get the Euclidean action

$$S_E(A) = -\frac{1}{2}\int_{\mathbb{R}}\int_{\mathbb{R}^3}\left((\partial_t A_0)^2 + ||\nabla A_0||^2\right) + \frac{1}{2}\sum_{j=1}^{3}\int_{\mathbb{R}}\int_{\mathbb{R}^3}\left((\partial_t A_j)^2 + ||\nabla A_j||^2\right),$$

and therefore we want to produce a probability measure with density proportional to $e^{-S_E(A)}$. Now there is a problem – the A_0 term ends up having a positive exponent – so the ad hoc way to fix this is that for the Euclidean free electromagnetic field, we have

$$A_0^E = i\phi_0, \quad A_j^E = \phi_j,$$

where $\phi_0, \phi_1, \phi_2, \phi_3$ are four independent Euclidean free fields on $\mathbb{R}^4$ with mass zero. So adding this factor of i formally gives us the right probability measure, and our components of the free electromagnetic field are just copies of the free field but with an extra i for the time component.

Our next task is to construct a Hilbert space and a semigroup; the Hilbert space is defined as follows. Define

$$\mathcal{U} = (i\mathbb{R})^{\mathcal{S}(\mathbb{R}^3)} \times \mathbb{R}^{\mathcal{S}(\mathbb{R}^3)} \times \mathbb{R}^{\mathcal{S}(\mathbb{R}^3)} \times \mathbb{R}^{\mathcal{S}(\mathbb{R}^3)},$$

endowed with the product of the four cylinder σ-algebras. Then a timeslice $(i\phi_0(0,\cdot), \phi_1(0,\cdot), \phi_2(0,\cdot), \phi_3(0,\cdot))$ is therefore a $\mathcal{U}$-valued random variable; we let μ be the law of $(i\phi_0, \phi_1, \phi_2, \phi_3)$ and $\mathcal{H} = L^2(\mu)$. We can then define the semigroup as a conditional expectation

$$P(t)F(X) = \mathbb{E}\left[F(A^E(t,\cdot))|A^E(0,\cdot) = X\right].$$

This is a strongly continuous Markov process on a semigroup – we can show this very similarly to how we did it with just a single copy of the free field. So this gives us a semigroup, and we can now try to work with and calculate with it – we'll calculate the photon propagator in the Feynman gauge (which is the analog of the Feynman propagator).

Definition 98

The **photon propagator** $\Pi = (\Pi_{j,k})_{0\le j,k\le 3}$ is defined by, for any $x, y \in \mathbb{R}^4$,

$$\Pi_{j,k}(x-y) = i\left\langle\Omega|\mathcal{T}(A_j(x)A_k(y))|\Omega\right\rangle,$$

where we now understand how to define each of these terms from our work before (in particular, $\mathcal{T}(A_j(x)A_k(y)) = A_j(x)A_k(y)$ if the first coordinate of x is smaller than that of y; otherwise it is $A_k(y)A_j(x)$).

Notice that this propagator is the expected value of something that is not gauge-invariant, hence this is not gauge invariant itself.

Proposition 99

For any $t \in \mathbb{R}$ and $x \in \mathbb{R}^3$, we have that

$$\Pi_{j,k}(t,x) = \frac{i\eta_{jk}}{(2\pi)^3}\int_{\mathbb{R}^3}\frac{e^{-i|t|\cdot||p||+ip\cdot x}}{2||p||}|dp,$$

where η is the matrix $\mathrm{diag}(-1,1,1,1)$. This expression is a distribution either for fixed t or on all four coordinates (the singularity here is benign).

Proof. Let $f, g \in \mathcal{S}(\mathbb{R}^3)$ be any Schwartz functions, and define $F_1(\phi) = \phi(g), F_2(\phi) = \phi(f)$. We'll take times $0, t$

just for simplicity; we want to compute

$$\langle \Omega | A_j(0,g) A_k(t,f) | \Omega \rangle ,$$

where by $A_j(0,g)$ we mean $F_1(A_j(0,\cdot) = \int_{\mathbb{R}^3} A_j(0,x) g(x) dx$ (and similar for the other term), so that our expression can also be written as $\langle \Omega | F_1(A_j(0,\cdot)) F_2(A_k(t,\cdot)) | \Omega \rangle$. This still doesn't quite make sense because we have quantum fields, but the way we rigorously defined this vacuum expectation (for a two-point correlation) is that we define a function

$$h(t) = \mathbb{E}\left[F_1(A_j^E(0,\cdot)) F_2(A_k^E(t,\cdot)) \right],$$

evaluate what this is on the nonnegative real line, and then extend by analytic continuation and continuity and finally evaluate on the imaginary line at it. So to do that computation, we can rewrite the above expectation in terms of our free fields as

$$h(t) = \mathbb{E}\left[F_1(\theta_j \phi_j(0,\cdot)) F_2(\theta_k \phi_k(t,\cdot)) \right], \quad \text{where } \theta = (i,1,1,1),$$

and by linearity this further simplifies to

$$\theta_j \theta_k \mathbb{E}\left[F_1(\phi_j(0,\cdot)) F_2(\phi_k(t,\cdot)) \right].$$

If j,k are different this is zero because we have independent Gaussians, and otherwise we've previously evaluated this (via Theorem 88 with mass $m=0$): we get that

$$\begin{aligned} h(t) &= \theta_j \theta_k \delta_{jk} \frac{1}{(2\pi)^3} \int \frac{e^{-t||p||} \hat{f}(p) \overline{\hat{g}(p)}}{2||p||} dp \\ &= \eta_{jk} \frac{1}{(2\pi)^3} \int \frac{e^{-t||p||} \hat{f}(p) \overline{\hat{g}(p)}}{2||p||} dp. \end{aligned}$$

So the Euclidean propagator is real-valued even with this weird definition with the i, and now we can indeed do analytic continuation with this expression because there's a minus in front of the t exponent. So we just replace t with it and that yields the desired result. □

Notice in particular that as $t, x \to 0$, the expression

$$\Pi_{j,k}(t,x) = \frac{i\eta_{jk}}{(2\pi)^3} \int_{\mathbb{R}^3} \frac{e^{-i|t| \cdot ||p|| + ip \cdot x}}{2||p|} |dp$$

blows up. So if t are small and g, f are close to delta functions that are close to each other, the calculation we have will get large.

We'll spend the next few lectures explaining "why photons" (why particles come in), skipping some proofs to get through the rest of the discussion.

26 March 10, 2025

Our topic for the last week of class is to show how the Hilbert space we've constructed is the same as the Fock space (describing the state space for particles, where the number of particles is indeterminate). The correspondence is not entirely clear at the moment, since we just have a quantum field.

Remark 100. *We won't have time to discuss this, but the Hilbert space for photons essentially has two copies rather than four. The reason for this is that the Hilbert space we constructed is not the physical one – the Lorentz gauge-fixing we did does not take away all degrees of freedom. For Schwartz functions $\Box\psi = 0$ implies $\psi = 0$, but in general we can have nonzero plane waves. And there's a certain quantization needed to quotient our Hilbert space out by*

linear maps. (But expectations of gauge-invariant observables still carry out in the way we did last time – it's just that we need to do a quotienting if we want to get to particles.)

To understand the thing that gets us from a state of the **free field** to bosons, we'll need to do a bit of discussion of special relativity. Consider $\mathbb{R}^4 = \mathbb{R}^{1,3}$ with the Minkowski inner product; we'll consider units where the speed of light is $c = 1$, so that the inner product reads

$$(x, y) = -x_0y_0 + x_1y_1 + x_2y_2 + x_3y_3.$$

Definition 101

A **Lorentz transform** on $\mathbb{R}^4$ is a linear map $L : \mathbb{R}^4 \to \mathbb{R}^4$ (that is, a 4×4 real matrix) such that $L^T\eta L = \eta$ for the Minkowski signature matrix $\eta = \text{diag}(-1, 1, 1, 1)$.

We may compare this with the identity for orthogonal matrices $L^TL = I$ – much like orthogonal matrices preserve the usual inner product, Lorentz transforms preserve the Minkowski inner product. These transformations form the group $O(1, 3)$.

Letting $O(4)$ denote the group of 4×4 orthogonal matrices and $SO(4)$ the subgroup of $O(4)$ with determinant 1, we have $SO(4) = \ker(\det)$, where $\det : O(4) \to \mathbb{Z}_2 = \{-1, 1\}$ is the determinant map. We can similarly go from $O(1, 3)$ to $SO(1, 3)$ by defining a map $\det : (1, 3) \to \mathbb{Z}_2$ (it's indeed valued in $\{\pm1\}$ because $\det(L)\det(\eta)\det(L) = \det(\eta)$) and defining $SO(1, 3) = \ker(\det)$.

However, it turns out we may further consider the group homomorphism $S(L) = \text{sgn}(L_{00})$; if we additionally quotient by the kernel of this, we get the **restricted Lorentz transformations**

$$SO^{\uparrow}(1, 3) = \{L \in O(1, 3) : \det(L) = 1, \text{sign}(L_{00}) > 0\}.$$

The key fact that we care about for this setting is that **under restricted Lorentz tarnsforms, physics remains invariant** (meaning that if we change coordinates, the result should remain the same; rephrased, the "allowable pictures" in these different coordinate systems are the same). Translations and Lorentz transforms combine to give the Poincaré group, which we won't talk about further.

Definition 102

Let x be the trajectory of a particle traveling through spacetime. We say that the trajectory is **parameterized by proper time** when written as $x(\tau)$ with $\left(\frac{dx}{d\tau}, \frac{dx}{d\tau}\right) = 1$.

This is the analogous construction to parameterizing by arc length in Euclidean space, which is invariant under transformations like rotations.

Example 103

The curve $(t, y(t))$ is parameterizable by proper time if and only if $||\frac{dy}{dt}|| < 1$ for all t.

In other words, unlike parameterization by arc length (which can basically always be done), parameterization by proper time can only be done if the speed of the particle is less than the speed of light c. And we care about doing this, because it gives us a coordinate-free notion of momentum:

Definition 104

Let $x(t)$ be a trajectory of a particle of mass m, parameterized by proper time τ. Then the **four-momentum** of the particle at any point on the trajectory is given by $p = m\frac{dx}{d\tau}$.

This is a coordinate-independent quantity, and in fact we can check that

$$p = \left(\frac{m}{\sqrt{1 - ||v(t)||^2}}, \frac{mv(t)}{\sqrt{1 - ||v(t)||^2}} \right),$$

where $v(t) = \frac{dy}{dt}$. Writing this in components as (p_0, p_1, p_2, p_3), we find that

$$-p_0^2 + p_1^2 + p_2^2 + p_3^2 = -\frac{m^2}{1 - ||v||^2} + \frac{m^2||v||^2}{1 - ||v||^2} = -m^2,$$

so $p^2 = -m^2$, and if we denote $q = (p_1, p_2, p_3)$ then we have the familiar factor $p_0 = \sqrt{m^2 + q^2} = \omega_q$. Now consider the following space (which has a Lorentz-invariant definition)

$$\mathcal{X}_m = \{p \in \mathbb{R}^4 : p^2 = -m^2, p_0 \geq 0\}.$$

In two dimensions this is the set of points (p_0, p_1) with $p_0 \geq 0$ and $-p_0^2 + p_1^2 = -m^2$ (which is one part of a hyperbola). In higher dimensions we get a similar shape, and we call this the **mass shell**. So the four-momentum of a particle is always in a particular mass-shell (and for $m = 0$ this is the light-cone for a photon).

Fact 105

Recall that in the nonrelativistic case, for a state $\psi \in L^2(\mathbb{R}^3)$, the probability density for the (three-vector) position is proportional to $|\psi(x)|^2$, and similarly the probability density for the (three-vector) momentum is proportional to $|\hat{\psi}(p)|^2$.

But in the relativistic (boson) case, position states are not very well-defined – instead, for relativistic scalar bosons of mass m, the four-momentum state states live in the space $L^2(\mathcal{X}_m, \lambda_m)$, where λ_m is the unique measure on $\mathcal{X}_m$ which is invariant under the restricted Lorentz transforms $SO^{\uparrow}(1, 3)$. (In the Euclidean case, this would be like having the uniform measure on the sphere – note that this is not exactly Haar measure on a group, since we have a group acting on the space instead of the group itself.)

To write things out more explicitly, there is a projection map $\pi : \mathcal{X}_M \to \mathbb{R}^3$ given by $\pi(p_0, p_1, p_2, p_3) = (p_1, p_2, p_3)$, and remember that by the mass shell condition we have $p_0 = \omega_q$.

Definition 106

For any m, define the measure λ_m on $\mathcal{X}_m$ by specifying that for any $A \in \mathcal{B}(\mathcal{X}_M)$,

$$\lambda_m(A) = \frac{1}{(2\pi)^3} \int_{\pi(A)} \frac{dq}{2\omega_q}.$$

In words, we're adding an additional weight to the usual Lebesgue measure on $\mathbb{R}^3$ to get invariance on the mass shell.

Lemma 107

Any restricted Lorentz transform $L \in SO^{\uparrow}(1, 3)$ is a bijection from $\mathcal{X}_m$ onto itself.

Proof. First we check that we do map into the space. For any $p \in \mathcal{X}_m$ and $L \in SO^{\uparrow}(1,3)$, let $q = Lp$. We then have

$$q^2 = q^T \eta q = p^T L^T \eta L p = p^T \eta p = p^2 = -m^2,$$

and we also need to check that $q_0 \geq 0$. But through a sequence of matrix computations, we see that

$$L^T \eta L = \eta \implies \eta L^T \eta L = \eta^2 = I \implies (\eta L)^{-1} = \eta L^T \implies \eta L \eta L^T = I \implies L \eta L^T = \eta,$$

and comparing the top-left elements of the last equality, we get

$$-L_{00}^2 + L_{01}^2 + L_{02}^2 + L_{03}^2 = -1 \implies \boxed{L_{01}^2 + L_{02}^2 + L_{03}^2} = L_{00}^2 - 1 < \boxed{L_{00}^2}.$$

Combined with the fact that $\boxed{p_1^2 + p_2^2 + p_3^2} = p_0^2 - m^2 \leq \boxed{p_0^2}$, we see that

$$\begin{aligned} q_0 &= L_{00}p_0 + L_{01}p_1 + L_{02}p_2 + L_{03}p_3 \\ &\geq L_{00}p_0 - \sqrt{(L_{01}^2 + L_{02}^2 + L_{03}^2)(p_1^2 + p_2^2 + p_3^2)} \\ &\geq 0, \end{aligned}$$

so we do indeed map into the mass-shell. The rest is easy from here: injectivity holds because the linear map L is injective, and for surjectivity note that L^{-1} is also a restricted Lorentz transform (we just need to check that $\text{sign}(L_{00}^{-1}) > 0$ and thus also maps from $\mathcal{X}_m$ into $\mathcal{X}_m$. □

What we'll do next is to "slightly expand" the mass shell

$$\mathcal{X}_{m,\varepsilon} = \bigcup_{m' \in [m, \sqrt{m^2+\varepsilon}]} \mathcal{X}_{m'} = \{p \in \mathbb{R}^4 : p^2 \in [-(m^2+\varepsilon), -m^2]\}.$$

We can then define the restriction of Lebesgue measure to this set

$$\lambda_{m,\varepsilon} = \frac{1}{(2\pi)^3 \varepsilon} \text{Leb}|_{\mathcal{X}_{m,\varepsilon}},$$

and since L is an injective linear map which maps $\mathcal{X}_{m,\varepsilon}$ to $\mathcal{X}_{m,\varepsilon}$, it also preserves $\lambda_{m,\varepsilon}$ (by the determinant formula for change of variable, noting that $\det(L) \in \{\pm 1\}$). We'll see that as we send $\varepsilon \to 0$, this goes down to a measure on the shell, and (with a bit of work) it will be exactly λ_m because the way we've projected down exactly gives us the extra factor in our measure. Then we'll have the (momentum) state space for a single relativistic boson, and we'll do the Fock space construction from there.

27 March 12, 2025

Our goal last time was to prove that the measure λ_m on the mass-shell is Lorentz-invariant. We started by expanding the mass shell to a thickening $\mathcal{X}_{m,\varepsilon} = \{p \in \mathbb{R}^4 : p^2 \in [-(m^2+\varepsilon), -m^2], p_0 \geq 0\}$ and defining $\lambda_{m,\varepsilon}$ to be Lebesgue measure restricted to that set; this is clearly Lorentz-invariant.

Lemma 108

For any $f \in \mathcal{S}(\mathbb{R}^4)$, we have

$$\lim_{\varepsilon \to 0} \int_{\mathcal{X}_{m,\varepsilon}} f(p) d\lambda_{m,\varepsilon}(p) = \int_{\mathcal{X}_m} f(p) d\lambda_m(p).$$

The subtlety here is that we need not just weak convergence $\lambda_{m,\varepsilon} \to \lambda_m$ but also a way to prove Lorentz invariance, which is why the claim is stated in this way.

Proof. Define $\omega_{q,\varepsilon} = \sqrt{\omega_q^2 + \varepsilon} = \sqrt{||q||^2 + m^2 + \varepsilon}$, then we have (no serious complications here since it's just Lebesgue integration)

$$\begin{aligned}\int_{\mathcal{X}_{m,\varepsilon}} f(p) d\lambda_{m,\varepsilon}(p) &= \frac{1}{(2\pi)^3 \varepsilon} \int_{\mathbb{R}^3} \int_{p_0:(p_0,q)\in X_{m,\varepsilon}} f(p_0, q) dp_0 dq \\ &= \frac{1}{(2\pi)^3 \varepsilon} \int_{\mathbb{R}^3} \int_{\omega_q}^{\omega_{q,\varepsilon}} f(p_0, q) dp_0 dq,\end{aligned}$$

and now we can apply the dominated convergence theorem (since f is Schwartz, so it's nice enough – just being bounded and continuous may cause problems). More precisely, we can write $f(p_0, q) = f(\omega_q, q) + O(\varepsilon)$ and the $O(\varepsilon)$ part will go away as $\varepsilon \to 0$ since we integrate in a window ε; thus this is roughly equal to

$$\frac{1}{(2\pi)^3 \varepsilon} \int_{\mathbb{R}^3} ((\omega_{q,\varepsilon} - \omega_q) f(\omega_q, q) dq) + O(\varepsilon),$$

where we can take the $O(\varepsilon)$ outside the integral because f has fast-decaying derivatives. So as $\varepsilon \to 0$, this converges to

$$\frac{1}{(2\pi)^3} \int_{\mathbb{R}^3} \frac{\omega_{q,\varepsilon} - \omega_q}{\varepsilon} f(\omega_q, q) dq,$$

and since $\omega_{q,\varepsilon} = \omega_q + \frac{\varepsilon}{2\omega_q} + O(\varepsilon^2)$, substituting this in indeed shows that we end up integrating on the projection down $\mathcal{X}_m$ and getting this weighting factor $\frac{1}{2\omega_p}$. □

So we have something nice in our hands now, and for Lorentz invariance we can now consider any $L \in SO^\uparrow(1,3)$ and write

$$\int_{\mathcal{X}_m} f(L(p)) d\lambda_m(p) = \lim_{\varepsilon \to 0} \int_{\mathcal{X}_{m,\varepsilon}} f(L(p)) d\lambda_{m,\varepsilon}(p)$$

since $f \circ L$ is also Schwartz. But now we know that $\lambda_{m,\varepsilon}$ is Lorentz invariant, so we can remove the L on the right-hand side and thus this is also equal to $\lim_{\varepsilon\to 0} \int_{\mathcal{X}_{m,\varepsilon}} f(p) d\lambda_{m,\varepsilon}(p) = \int_{\mathcal{X}_m} f(p) d\lambda_m(p)$. This is for the case $f \in \mathcal{S}(\mathbb{R}^4)$, though, and now we want to consider Schwartz functions **on the shell** rather than Schwartz functions on all of $\mathbb{R}^4$.

Definition 109

Let $\pi : \mathcal{X}_m \to \mathbb{R}^3$ be the projection onto the spatial coordinates. We say that a function $h : \mathcal{X}_m \to \mathbb{R}$ is **Schwartz** if $h = g \circ \pi$ for some $g \in \mathcal{S}(\mathbb{R}^3)$.

One perspective for this is that if we think of $\mathcal{X}_m$ as a 3-dimensional manifold, then this projection map gives us a coordinate chart, and we can carry over definitions (such as smoothness) from the chart to the manifold. And now our **goal** is to show that

$$\int_{\mathcal{X}_m} h(L(p)) d\lambda_m(p) = \int_{\mathcal{X}_m} h(p) d\lambda_m(p)$$

for Schwartz functions $h : \mathcal{X}_m \to \mathbb{R}$. Then Lorentz invariance will follow (that is, $\lambda_m(L^{-1}(A)) = \lambda_m(A)$ for all Borel sets $A \in \mathcal{B}(\mathcal{X}_m)$, because indicator functions can be approximated by Schwartz functions. For this, let $g \in \mathcal{S}(\mathbb{R}^3)$ be the corresponding function such that $h = g \circ \pi$, and define $f : \mathbb{R}^4 \to \mathbb{R}$ via

$$f(p_0, q) = g(q) e^{-(p_0^2 - \omega_q - m^2)^2}.$$

This is Schwartz on $\mathbb{R}^4$ and coincides with h on $\mathcal{X}_m$, so our result follows.

The proof that this is the unique invariant measure will be omitted, but the point is that up to constant multiples this is the correct one on the mass shell.

Recall now that our goal is to show that the state space for the quantum free field is the same as that for an indeterminate number of bosons. So we'll first represent the former in a slightly different way from how we've been doing it so far – instead of the L^2 space of a bunch of Gaussian variables, what we really need is the Fourier transform of this field (which is exactly the same, but reindexed in a different way). Formally, the Fourier transform is given by (for $q \in \mathbb{R}^3$)

$$\hat{\phi}_0(q) = \int_{\mathbb{R}^3} \phi_0(x) e^{iq\cdot x} dx;$$

of course, this doesn't make sense pointwise because we can only integrate against Schwartz functions, not things like $e^{iq\cdot x}$. But note that ϕ_0 is real-valued (meaning it's real-valued when integrated against Schwartz functions), but $\hat{\phi}_0$ should be complex-valued. So we can think of it as two fields, one for the real part and one for the complex part:

Definition 110

For $f \in \mathcal{S}(\mathbb{R}^3)$, define

$$\mathrm{Re}\,\hat{\phi}_0(f) = \phi_0(\mathrm{Re}(\hat{f})), \quad \mathrm{Im}\,\hat{\phi}_0(f) = -\phi_0(\mathrm{Im}(\hat{f})).$$

The (formal, heuristic) justification for this definition is that if $g(x) = \hat{f}(-x)$, then by Fourier inversion

$$\begin{aligned} f(q) &= \frac{1}{(2\pi)^3} \int_{\mathbb{R}^3} e^{-ix\cdot q} \hat{f}(x) dx \\ &= \frac{1}{(2\pi)^3} \int_{\mathbb{R}^3} e^{ix\cdot q} \hat{f}(-x) dx \\ &= \frac{1}{(2\pi)^3} \int_{\mathbb{R}^3} e^{ix\cdot q} g(x) dx, \end{aligned}$$

and this last expression is exactly $\frac{1}{(2\pi)^3}\hat{g}(q)$. Therefore (in the second equality we use that f is real-valued)

$$\begin{aligned} \hat{\phi}_0(f) &= \int \hat{\phi}_0(q) f(q) dq \\ &= \int \hat{\phi}_0(q) \overline{f(q)} dq \\ &= \frac{1}{(2\pi)^3} \int \hat{\phi}_0(q) \overline{\hat{g}(q)} dq, \end{aligned}$$

and by Plancherel this is $\int \phi_0(x) g(x) dx = \phi_0(\mathrm{Re}(g)) + i\phi_0(\mathrm{Im}(g))$. So those two expressions should be the real and imaginary parts of $\hat{\phi}_0(f)$, and it just remains to figure out what $\mathrm{Re}(g)$ and $\mathrm{Im}(g)$ are. But

$$\begin{aligned} g(x) = \hat{f}(-x) &= \int e^{-ix\cdot q} f(q) dq \\ &= \overline{\int e^{ix\cdot q} f(q) dq} \\ &= \overline{\hat{f}(x)} \\ &= \mathrm{Re}(\hat{f}(x)) - i\mathrm{Im}(\hat{f}(x)), \end{aligned}$$

and therefore $\mathrm{Re}(g) = \mathrm{Re}(\hat{f})$ and $\mathrm{Im}(g) = -\mathrm{Im}(\hat{f})$, justifying the definition we just made.

We can now calculate the covariance structure in this "re-indexed way:"

Lemma 111

Let $\omega_q = \sqrt{||q||^2 + m^2}$ as usual. For any Schwartz functions $f, g \in \mathcal{S}(\mathbb{R}^3)$,

$$\text{Cov}\left(\text{Re}\,\hat{\phi}_0(f), \text{Im}\,\hat{\phi}_0(g)\right) = 0,$$

$$\text{Cov}\left(\text{Re}\,\hat{\phi}_0(f), \text{Re}\,\hat{\phi}_0(g)\right) = \frac{(2\pi)^3}{4}\int_{\mathbb{R}^3} \frac{1}{2\omega_q}(f(q) + f(-q))(g(q) + g(-q))dq.$$

$$\text{Cov}\left(\text{Im}\,\hat{\phi}_0(f), \text{Im}\,\hat{\phi}_0(g)\right) = \frac{(2\pi)^3}{4}\int_{\mathbb{R}^3} \frac{1}{2\omega_q}(f(q) - f(-q))(g(q) - g(-q))dq.$$

(Notice that we're taking the Green's function restricted to the timeslice $t = 0$, which is why we have ω_q instead of ω_q^2.) These expressions are quite different from how ϕ_0 involved f and g at very distant points, while this covariance structure is "almost diagonal." One way to say this is that it's "like white noise" where we have independent values at different points, except that the variables at q and $-q$ are the same or negatives of each other for real or imaginary, respectively – in other words, $\hat{\phi}_0(q)$ and $\hat{\phi}_0(-q)$ are complex conjugates of each other.

We won't do the full proof here; just recall that we computed the covariance structure in terms of the Fourier transform while computing the Feynman propagator, so we just write down the definitions such as

$$\text{Cov}\left(\text{Re}\,\hat{\phi}_0(f), \text{Im}\,\hat{\phi}_0(g)\right) = \text{Cov}\left(\phi_0(\text{Re}(\hat{f}), -\phi_0(\text{Im}(\hat{g})\right)$$

and plug in our known values.

What we will do now is transfer things to the mass shell (that is, from $\mathbb{R}^3$ to $\mathcal{X}_m$). For any $f \in \mathcal{S}(\mathcal{X}_m)$, meaning that $f = g \circ \pi$ for some unique $g \in \mathcal{S}(\mathbb{R}^3)$, we can define the complex-valued field

$$\psi(f) = \text{Re}\,\hat{\phi}_0(g) + i\text{Im}\,\hat{\phi}_0(g) = \hat{\phi}_0(g)$$

where we write $\text{Re}\,\hat{\phi}_0(g)$ as $\psi_1(f)$ and similarly write $\text{Im}\,\hat{\phi}_0(g)$ as $\psi_2(f)$.

Recall that the Hilbert space $\mathcal{H}$ for the free quantum field is the L^2 space of the random variables $\{\phi_0(f) : f \in \mathcal{S}(\mathbb{R}^3)\}$. We have the useful approximation property (which we will not prove) that for any n and any Gaussian measure γ on $\mathbb{R}^n$, polynomials are dense in $L^2(\gamma)$. Also recall that (via measure-theoretic arguments) functions of finitely many $\phi_0(f)$s are dense in our Hilbert space. Thus, **polynomials in $\phi_0(f)$s are dense in our Hilbert space**, because any L^2 function can be approximated by one of finitely many coordinates, and then we can approximate by polynomials. This means also that letting ψ be the complex-valued field we just defined, **polynomials in $\psi_1(f), \psi_2(f)$, where f ranges over all Schwartz functions $\mathcal{S}(\mathcal{X}_m)$, are dense in $\mathcal{H}$ as well.**

It turns out we can go from real and imaginary-valued functions to just single objects:

Lemma 112

Polynomials in $\{\psi(f) : f \in \mathcal{S}(\mathcal{X}_m, \mathbb{C})\}$ are dense in $\mathcal{H}$.

This is not immediate: we have to show that the complex conjugate of any function like this is equal to ψ of some other function, and that's not true for all fields but it's true here "formally because $\psi(q) = \overline{\psi(-q)}$."

Proof sketch. We have

$$\text{Re}\,\psi(f) = \psi_1(f) = \frac{1}{2}(\psi(f) + \overline{\psi(f)}), \quad \text{Im}\,\psi(f) = \psi_2(f) = \frac{1}{2i}(\psi(f) - \overline{\psi(f)}),$$

we do indeed only need to prove that $\overline{\psi(f)} = \psi(g)$ for some g. Take any complex-valued Schwartz function on the mass shell $f \in \mathcal{S}(\mathcal{X}_m, \mathbb{C})$; we can write any point $p \in \mathcal{X}_m$ as (ω_q, q) for some unique $q \in \mathbb{R}^3$, and we define $\tilde{p} = (\omega_q, -q)$. If we then define $g(p) = \overline{f(\tilde{p})}$, we can show using the covariance structure of ψ above that

$$\mathbb{E}\left[\left|\overline{\psi(f)} - \psi(g)\right|^2\right] = 0.$$

□

With all of this work, we've expressed our Hilbert space as a closure of polynomials in $\psi(f)$ (as f ranges over all complex-valued Schwartz functions). We'll then break them into orthogonal subspaces in an explicit way, and this will give us the isomorphism required.

28 March 14, 2025

Recall the notation of the mass shell $\mathcal{X}_m$ and the measure λ_m we've discussed in the last few lectures; we'll now use this to construct the particle space. For a single scalar boson, the momentum state space is $L^2(\mathcal{X}_m, \lambda_m)$. We then say that for any state $\psi \in L^2(\mathcal{X}_m, \lambda_m)$, the probability of its 4-momentum lying in a Borel set $A \in \mathcal{B}(\mathcal{X}_m)$ is

$$\mathbb{P}(\text{4-momentum} \in A) = \frac{\int_A |\psi(p)|^2 d\lambda_m}{\int_{\mathcal{X}_m} |\psi(p)|^2 d\lambda_m}.$$

Now if we have n bosons of mass m, we may think that the state space will be $L^2(\mathcal{X}_m^n, \lambda_m^n)$, but in fact we should be considering the **symmetric subspace**

$$\begin{aligned} L^2_{\text{sym}}(\mathcal{X}_m^n, \lambda_m^n) &= \left\{f \in L^2(\mathcal{X}_m^n, \lambda_m^n) : f \text{ symmetric}\right\} \\ &= \left\{f \in L^2(\mathcal{X}_m^n, \lambda_m^n) : f(p_1, \cdots, p_n) = f(p_{\sigma(1)}, \cdots, p_{\sigma(n)}) \ \lambda_m^n\text{-almost-everywhere for all } \sigma \in S_n\right\}; \end{aligned}$$

that is, permuting the coordinates does not cause f to change. So wavefunctions can only be symmetric, meaning that the particles are in some sense indistinguishable.

But as we have previously discussed, the number of particles may be arbitrary as well, and thus we will also construct the **Fock space** for a system of bosons with mass m

$$\mathcal{B} = \widehat{\oplus_{n\geq 0}} \mathcal{B}_n,$$

where $\mathcal{B}_n = L^2_{\text{sym}}(\mathcal{X}_m^n, \lambda_m^n)$ is the symmetrized space we just constructed (and $\mathcal{B}_0$ is the complex line), and $\widehat{\oplus_{n\geq 0}}$ denotes the completion of the direct sum of $\mathcal{B}_0, \mathcal{B}_1, \mathcal{B}_2$, and so on. In other words,

$$\mathcal{B} = \left\{(\psi_0, \psi_1, \psi_2, \cdots) : \psi_n \in \mathcal{B}_n \text{ for all } n, \quad \sum_{n=0}^{\infty} ||\psi_n||^2 < \infty\right\}$$

is the set of sequences of wavefunctions where the sum of squared norms is finite. We then define the completed inner product

$$\langle (\psi_n)_{n\geq 0}, (\phi_n)_{n\geq 0} \rangle = \sum_{n=0}^{\infty} \langle \psi_n, \phi_n \rangle,$$

and in particular this makes all of the $\mathcal{B}_n$s orthogonal to each other. So the interpretation of this is that a state

$\psi = (\psi_n)_{n\geq 0}$ written as a sequence of wavefunctions satisfies

$$\mathbb{P}(\text{number of particles} = n) = \frac{||\psi_n||^2}{\sum_{j=0}^{\infty} ||\psi_j||^2},$$

and if we condition on the number of particles being n, the joint density of the 4-momenta of those n particles is proportional to $|\psi_n|^2$. (Note that we now still have "rays" in the Hilbert space, but we have to scale by the same constant in each $\mathcal{B}_n$ rather than doing so independently.)

We'll now correspond elements of this Fock space to states of the free quantum field. For this, we'll make use of the following:

Theorem 113 (Wick product formula)
Let $X_1, \cdots, X_n$ be jointly centered Gaussian random variables. Then letting $\mathcal{P}(\{1, \cdots, n\})$ be the set of **perfect matchings** of $\{1, \cdots, n\}$ (that is, the set of ways to group the elements into pairs, or equivalently simple undirected graphs where all vertices have degree 1), we have

$$\mathbb{E}[X_1 \cdots X_n] = \sum_{E \in \mathcal{P}(\{1,\cdots,n\})} \prod_{\{i,j\} \in E} \mathbb{E}[X_i X_j]$$

For example, regardless of the covariances between the random variables,

$$\mathbb{E}[X_1X_2X_3X_4] = \mathbb{E}[X_1X_2]\mathbb{E}[X_3X_4] + \mathbb{E}[X_1X_3]\mathbb{E}[X_2X_4] + \mathbb{E}[X_1X_4]\mathbb{E}[X_2X_3].$$

but on the other hand we always have $\mathbb{E}[X_1X_2X_3] = 0$.

Proof. We can prove this by integration by parts – for all well-behaved functions (in particular polynomials f), by integration by parts we have for Z standard normal that

$$\mathbb{E}[Zf(Z)] = \mathbb{E}[f'(Z)],$$

so if $Z_1, \cdots, Z_n$ are iid standard normal we also have (by Fubini's theorem and then applying the above result)

$$\mathbb{E}[Z_j f(Z_1, \cdots, Z_n)] = \mathbb{E}[\partial_j f(Z_1, \cdots, Z_n)].$$

So now if $X_1, \cdots, X_n$ are correlated centered Gaussians with covariance matrix Σ, we can find some matrix A with $\Sigma = AA^T$, and then we have $X = AZ$ for Z an n-dimensional standard Gaussian. We then find that

$$\mathbb{E}[X_j f(X)] = \sum_{k=1}^{n} \mathbb{E}[X_j X_k]\, \mathbb{E}[\partial_k f(X)],$$

and the Wick product formula now follows by iterating this identity: letting f be everything except X_1, we get

$$\mathbb{E}[X_1X_2 \cdots X_n] = \sum_{j=2}^{n} \mathbb{E}[X_1X_j]\mathbb{E}\left[\prod_{\ell \neq 1,k} X_\ell\right],$$

and induction does the rest from here (since 1 is matched with an arbitrary j here). □

We can now generalize to the complex Gaussian case – here a **complex Gaussian** random variable is a $\mathbb{C}$-valued random variable whose real and imaginary parts are jointly Gaussian.

Corollary 114

The Wick product formula holds in the exact same form for centered complex Gaussian random variables as well (by dividing variables up into real and complex parts).

Definition 115

Let $\{X_i\}_{i\in I}$ be a (possibly infinite) collection of jointly complex Gaussian random variables with mean zero (such as the field that we constructed last time). For any $i_1, \cdots, i_n \in I$ we define the **Wick ordered monomial** (let $V = \{1, \cdots, n\}$)

$$:X_{i_1}\cdots X_{i_n}: = \sum_{\substack{U\subseteq V\\ |U| \text{even}}} (-1)^{|U|/2} \left(\sum_{E\in\mathcal{P}(U)} \prod_{\{a,b\}\in E} \mathbb{E}\left[X_{i_a}X_{i_b}\right]\right) \prod_{a\in V\setminus U} X_{i_a}.$$

The leading term here comes from U being the emptyset, which just yields $X_{i_1}\cdots X_{i_n}$, but we can also replace some even-sized subset by a Wick-type term. We can see for example that

$$:X_1X_2X_3: = X_1X_2X_3 - \mathbb{E}[X_1X_2]X_3 - \mathbb{E}[X_1X_3]X_2 - \mathbb{E}[X_2X_3]X_1,$$

which is some polynomial in the X_is. In particular taking all indices to be the same, we get

$$:X^3: = X^3 - 3X,$$

which we may notice is the Hermite polynomial of degree 3. This isn't a coincidence, and the main remarkable property of such objects is the following:

Theorem 116

Every Wick-ordered monomial has zero expected value, and the product of two Wick-ordered monomials

$$\mathbb{E}\left[:X_{i_1}\cdots X_{i_n}: :X_{j_1}\cdots X_{i_m}:\right] = \begin{cases} \sum_{\sigma\in S_n}\prod_{a=1}^n \mathbb{E}\left[X_{i_a}X_{j_{\sigma(a)}}\right] & m = n, \\ 0 & \text{otherwise.}\end{cases}$$

We'll skip the proof – it basically comes down to applying the Wick formula and doing some pretty messy calculations. So these are kind of the generalizations of the Hermite polynomials for a single Gaussian – those polynomials are chosen so that they generate orthogonal subspaces, and here an arbitrary (even infinite) collection of variables we can do that same decomposition.

Turning back to the field ψ that we previously defined, recall that the way we constructed it is to take the Euclidean free field ϕ_0 at time 0, consider its Fourier transform $\hat{\phi}_0$, and then "lift it" to ϕ on the mass shell $\mathcal{X}_m$. More precisely, for any complex-valued Schwartz function on the mass shell $f \in \mathcal{S}(\mathcal{X}_m, \mathbb{C})$, we have $f = g\circ\pi$ for some $g \in \mathcal{S}(\mathbb{R}^3, \mathbb{C})$, and then we defined

$$\psi(f) = \hat{\phi}_0(g) = \hat{\phi}_0(\text{Re}(g)) + i\hat{\phi}_0(\text{Im}(g)).$$

The Hilbert space for this free quantum field is then $\mathcal{H} = L^2(\psi)$ (this is shorthand for saying that ψ is a collection of infinite Gaussians, we induce some probability measure on that space, and then we take the L^2 space of that); we can now construct an explicit Hilbert space isomorphism between $\mathcal{H}$ and the bosonic Fock space $\mathcal{B}$. We define

$$\mathcal{H}_n = \text{subspace of } \mathcal{H} \text{ spanned by Wick-ordered monomials of } \{\psi(f) : f \in \mathcal{S}(\mathcal{X}_m, \mathbb{C})\} \text{ of degree } n,$$

and define the map $\tau : \bigoplus \mathcal{H}_n \to \mathcal{B}$ via

$$\tau\left(:\psi(f_1)\cdots\psi(f_n):\right) = F, \text{ where } F(p_1, \cdots, p_n) = \frac{1}{\sqrt{n!}} \sum_{\sigma \in S_n} f_1(p_{\sigma(1)}) \cdots f_n(p_{\sigma(n)}).$$

The right-hand side is a symmetric function of $p_1, \cdots, p_n$, so it indeed lies in $\mathcal{B}_n$. We then extend τ by linearity of $\mathcal{H}_n$ and then to the full space $\mathcal{H}$; it's not quite clear from what we said that this is well-defined, but it is. It turns out τ preserves inner products (this is the reason for the $\frac{1}{\sqrt{n!}}$ factor, and here is where we use the uncorrelated nature of the Wick monomials), hence is an isometry, and thus extends to the closure of $\bigoplus_{n \geq 0} \mathcal{H}_n$, which is all of $\mathcal{H}$ (since we can recover all polynomials from the Wick monomials, and polynomials are dense in Gaussian space). Furthermore, τ is surjective – any function can be approximated in L^2 by functions which are products of coordinate functions.

Playing around with this, we can thus see "what the field looks like." For example, if we have a single particle localized in a region and moving at some speed, we can get a corresponding state in the field and study its properties that way.