Quantum Fields and Particles

Felix Flicker

7th August 2025

Contents

List of Definitions assumed from Quantum Mechanics

The canonical commutation relation $[\hat{x}, \hat{p}] = i\hbar\hat{\mathbb{I}}$

Dirac notation the notation $|\psi\rangle$ for complex vectors. Also called bra-ket notation, with $\langle\phi|$ the 'bra', $|\psi\rangle$ the 'ket', and $\langle\phi|\psi\rangle$ a bracket.

Expectation value $\langle\hat{A}\rangle = \langle\psi|\hat{A}|\psi\rangle$. The mean value of an operator measured by a given state.

First quantization a wave-like description of quantum objects: $\psi(x)$.

The Hamiltonian the energy operator (assumed time independent in this course). $\hat{H} = \hat{p}^2/2m + \hat{V}$, or $\hat{H}\psi(x) = -\hbar^2\psi''/2m + V(x)\psi$.

The Heisenberg picture the description of quantum states as time independent, and operators as time dependent.

The Heisenberg uncertainty principle $\sigma_{\hat{A}}\sigma_{\hat{B}} \geq \frac{1}{2}\left|\left\langle\left[\hat{A}, \hat{B}\right]\right\rangle\right|$ where $\sigma_{\hat{A}}$ denotes the standard deviation of operator $\hat{A}$.

Hilbert space a linear vector space with an inner product and square-normalisable vectors

Hermiticity $\hat{A} = \hat{A}^\dagger$ where $\hat{A}^\dagger = \hat{A}^{*T}$. For differential operators: $\int_{-\infty}^{\infty}\varphi(x)^*\left(\hat{A}\psi(x)\right)\mathrm{d}x = \int_{-\infty}^{\infty}\left(\hat{A}\varphi(x)\right)^*\psi(x)\,\mathrm{d}x$.

Ladder operators an operator which raises or lowers the quantum number of a state it acts on. Also called creation/annihilation operators or raising and lowering operators.

Normalisation the prefactor on a wavefunction ensuring that the total probability to find the particle is one.

The number operator in the harmonic oscillator, the operator whose eigenstates are the energy eigenstates and whose eigenvalues are the level of the state.

Operators objects which act on states to give states (either the same state, or a different one). In fininte dimensional Hilbert spaces these are simply matrices acting on vectors to give vectors. In infinite dimensional Hilbert spaces these are differential operators acting on functions to give functions.

Orthonormality orthogonal and normalised. If a set of states is orthonormal the inner product of any state with itself is 1 and the inner product between any two different states is zero.

The probability density ρ integrated over a region of space, this gives the probability to find the particle in that region. $\rho(x,t)\,\mathrm{d}x = |\psi(x,t)|^2\,\mathrm{d}x$ is the probability to find the particle between x and $x + \mathrm{d}x$ at time t.

The probability current density j the current density associated with a flow of probability: $\boldsymbol{j}(\boldsymbol{x},t) = \frac{i\hbar}{2m}\{\psi\nabla\psi^* - \psi^*\nabla\psi\}$.

The probability amplitude the complex number associated to each point in space by the wavefunction ψ.

Quantum numbers eigenvalues of operators which commute with the Hamiltonian; expectation values which do not change in time.

The Schrodinger picture the description of quantum states as time dependent, and operators as time independent.

Second quantisation a particle-like description of quantum objects in terms of ladder operators.

Stationary states energy eigenstates. So called as their probability densities are time independent.

Superposition summing solutions to the TDSE to get a new solution to the TDSE

The time dependent schroedinger equation $i\hbar\partial_t|\psi\rangle = \hat{H}|\psi\rangle$, or in the position basis $i\hbar\dot{\psi} = \hat{H}\psi$. Abbreviated TDSE.

The time independent schroedinger equation $\hat{H}|\psi\rangle = E|\psi\rangle$, or in the position basis $\hat{H}\psi(x) = E\psi(x)$. Abbreviated TISE.

The wavefunction ψ a function which assigns a complex number to each point in space. The modulus square is the probability density ρ.

Conventions on the Fourier transform

You will recall there are various conventions adopted with regard to factors of 2π in Fourier transforms. To my mind the simplest to remember is

$$\tilde{f}(\boldsymbol{k}) = \int \mathrm{d}^D \boldsymbol{x} \exp(-2\pi i \boldsymbol{k}\cdot\boldsymbol{x}) f(\boldsymbol{x}) \tag{1}$$

$$f(\boldsymbol{x}) = \int \mathrm{d}^D \boldsymbol{k} \exp(2\pi i \boldsymbol{k}\cdot\boldsymbol{x}) \tilde{f}(\boldsymbol{k}). \tag{2}$$

From here you can rescale variables as you like, for example:

$$\boldsymbol{k} \to \boldsymbol{k}/2\pi \tag{3}$$

gives

$$\tilde{f}(\boldsymbol{k}) = \int \mathrm{d}^D \boldsymbol{x} \exp(-i \boldsymbol{k}\cdot\boldsymbol{x}) f(\boldsymbol{x}) \tag{4}$$

$$f(\boldsymbol{x}) = \int \frac{\mathrm{d}^D \boldsymbol{k}}{(2\pi)^D} \exp(i \boldsymbol{k}\cdot\boldsymbol{x}) \tilde{f}(\boldsymbol{k}) \tag{5}$$

(where $\tilde{f}(\boldsymbol{k})$ has a rescaled argument).
Since I will use the $(+---)$ signature, spacetime Fourier transforms in $D+1$ dimensions must have the opposite signs in the exponents:

$$\tilde{f}(k) = \int \mathrm{d}^{D+1} x \exp(2\pi i k_\mu x^\mu) f(x) \tag{6}$$

$$f(x) = \int \mathrm{d}^{D+1} k \exp(-2\pi i k_\mu x^\mu) \tilde{f}(k). \tag{7}$$

For the discrete Fourier transform I will use

$$\tilde{f}_k = \sum_{n=1}^{N} \exp(-2\pi i k n/N) f_n \tag{8}$$

$$f_n = \frac{1}{N} \sum_{k=1}^{N} \exp(2\pi i k n/N) \tilde{f}_k. \tag{9}$$

Units and Dimensions

These notes will adopt the common practice of using 'natural units' with $\hbar = c = 1$. This sets energy/angular frequency ($E = \hbar\omega = pc$) and space/time ($ct = x$) on equal footing.

1 Introduction & Background

1.1 Lagrangian Classical Mechanics

In Lagrangian mechanics we work with generalised co-ordinates[1]

$$x_n(t) \triangleq x_{n,t} \tag{10}$$

and generalised velocities

$$\dot{x}_n(t) \triangleq \dot{x}_{n,t}$$

where $n \in [1, N]$. I will typically neglect to write the explicit time dependence of x_n and $\dot{x}_n$ when it is clear. For example, a ball moving in 3D would require $N = 3$ components of position, and three of velocity, to specify its state.
Define the Lagrangian

$$L = T - V \tag{11}$$

where T is the kinetic energy and V the potential energy. Specifically, in the absence of magnetic fields,

$$L = \sum_{n=1}^{N} \frac{1}{2} m \dot{x}_n^2 - V(x_n). \tag{12}$$

It will be convenient to define the function which is being summed over:

$$L = \sum_{n=1}^{N} \mathcal{L}(x_n, \dot{x}_n, t). \tag{13}$$

When we move to fields, $\mathcal{L}$ is called the *Lagrangian density*. It can include an explicit dependence on time (via a time-varying potential). We will typically not consider such cases, just as we typically don't include an explicit time dependence in Hamiltonians in quantum mechanics. Hence, time dependence will typically enter only implicitly via $x_{n,t}$ and $\dot{x}_{n,t}$.
Define the time integral of the Lagrangian to be

the *action* S:

$$S[\boldsymbol{x}] = \int_{t_0}^{t_f} L \mathrm{d}t = \int_{t_0}^{t_f} \sum_{n=1}^{N} \mathcal{L}(x_n, \dot{x}_n)\, \mathrm{d}t \tag{14}$$

where

$$[\boldsymbol{x}]_n = x_n. \tag{15}$$

In Lagrangian mechanics, we speciy the problem by writing down the action, then we find the dynamics using

the **principle of least action**:

classical paths $\boldsymbol{x}(t)$ extremise the action S.

Of course, in classical mechanics, all paths are classical, so these are simply the solutions to the problem you'd like to solve. When we move to quantum mechanics, there will be amplitudes associated with non-classical paths.
To extremise the action we use:

$$\left(\frac{\partial S[\boldsymbol{x} + \lambda \boldsymbol{\epsilon}]}{\partial \lambda} \right)_{\boldsymbol{x}, \boldsymbol{\epsilon}} \Bigg|_{\lambda = 0} = 0. \tag{16}$$

[1] It is common in classical mechanics to use the symbol q_i instead of x_n, but I will avoid this as it is much more common to use x_n in quantum mechanics.

Here, the subscripts to the parentheses indicate what is held constant in the partial derivative (always good practice!). Then λ is set to 0 after the differentiation. $\boldsymbol{\epsilon}(t)$ parameterises a variation away from the classical path $\boldsymbol{x}(t)$. We require the variation to be zero at the start and end points of the trajectory in order for the boundary conditions to be obeyed:

$$\boldsymbol{\epsilon}(t_0) = \boldsymbol{\epsilon}(t_f) = 0. \tag{17}$$

In the general case we find

$$S[\boldsymbol{x} + \lambda\boldsymbol{\epsilon}] = \int_{t_0}^{t_f} \sum_n \mathcal{L}(x_n + \lambda\epsilon_n, \dot{x}_n + \lambda\dot{\epsilon}_n)\,\mathrm{d}t \tag{18}$$

$$\Downarrow \text{ chain rule}$$

$$\left(\frac{\partial S[\boldsymbol{x} + \lambda\boldsymbol{\epsilon}]}{\partial\lambda}\right)_{\boldsymbol{x},\boldsymbol{\epsilon}} = \int_{t_0}^{t_f} \left\{\sum_n \left(\frac{\partial\mathcal{L}}{\partial x_n}\right)_{\dot{x}_n} \left(\frac{\partial(x_n + \lambda\epsilon_n)}{\partial\lambda}\right)_{x_n,\epsilon_n} + \left(\frac{\partial\mathcal{L}}{\partial\dot{x}_n}\right)_{x_n} \left(\frac{\partial(\dot{x}_n + \lambda\dot{\epsilon}_n)}{\partial\lambda}\right)_{\dot{x}_n,\dot{\epsilon}_n}\right\}\mathrm{d}t \tag{19}$$

$$= \int_{t_0}^{t_f} \sum_n \left\{\left(\frac{\partial\mathcal{L}}{\partial x_n}\right)_{\dot{x}_n} \epsilon_n + \left(\frac{\partial\mathcal{L}}{\partial\dot{x}_n}\right)_{x_n} \dot{\epsilon}_n\right\}\mathrm{d}t. \tag{20}$$

Now integrate the $\dot{\epsilon}_n$ term by parts in Eq 20:

$$\left(\frac{\partial S[\boldsymbol{x} + \lambda\boldsymbol{\epsilon}]}{\partial\lambda}\right)_{\boldsymbol{x},\boldsymbol{\epsilon}} = \int_{t_0}^{t_f} \sum_n \left\{\left(\frac{\partial\mathcal{L}}{\partial x_n}\right)_{\dot{x}_n} - \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial\mathcal{L}}{\partial\dot{x}_n}\right)_{x_n}\right\}\epsilon_n\mathrm{d}t + \left[\sum_n \left(\frac{\partial\mathcal{L}}{\partial\dot{x}_n}\right)_{x_n} \epsilon_n\right]_{t_0}^{t_f} \tag{21}$$

but the boundary term vanishes by assumption (Eq 17). Applying the principle of least action, Eq 16, we require

$$\left(\frac{\partial S[\boldsymbol{x} + \lambda\boldsymbol{\epsilon}]}{\partial\lambda}\right)_{\boldsymbol{x},\boldsymbol{\epsilon}}\Bigg|_{\lambda=0} = 0 = \int_{t_0}^{t_f} \sum_n \left\{\left(\frac{\partial\mathcal{L}}{\partial x_n}\right)_{\dot{x}_n} - \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial\mathcal{L}}{\partial\dot{x}_n}\right)_{x_n}\right\}\epsilon_n\mathrm{d}t. \tag{22}$$

This is true for all $\epsilon_n(t)$ (since this arbitrary function has not been specified). Therefore we have the **Euler Lagrange equations**:

$$\left(\frac{\partial\mathcal{L}}{\partial x_n}\right)_{\dot{x}_n} - \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial\mathcal{L}}{\partial\dot{x}_n}\right)_{x_n} = 0. \tag{23}$$

You will more commonly see this written in terms of the Lagrangian as

$$\left(\frac{\partial L}{\partial\boldsymbol{x}}\right)_{\dot{\boldsymbol{x}}} - \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial\dot{\boldsymbol{x}}}\right)_{\boldsymbol{x}} = 0 \tag{24}$$

where the derivative with respect to a vector is understoof component-wise as in Eq 23. With the general Lagrangian of Eq 12 we arrive at Newton's 2nd law:

$$-V'(\boldsymbol{x}) = m\ddot{\boldsymbol{x}}. \tag{25}$$

1.2 Functionals

The action is a *functional*. In general, a function takes in a number and returns a number. A functional instead takes in a function and returns a number.
Specifically, the action takes as its argument the *entire path* $x_{n,t}$. To see this, note that the action involves an integral over t and a sum over n. Hence, it cannot itself depend on t or n. It returns a number S.
In contrast, the Lagrangian density is a function. It takes in a specific instance of $x_{n,t}$ and $\dot{x}_{n,t}$ and returns a number $\mathcal{L}$.
A functional is written using square brackets $S[\boldsymbol{x}]$ as opposed to the parentheses of a function $\mathcal{L}(x_{n,t}, \dot{x}_{n,t})$.

We can derive the Euler Lagrange equations more succinctly by defining

the *functional derivative*

$$\frac{\delta S\left[\boldsymbol{x}\right]}{\delta \boldsymbol{x}} \triangleq \left(\frac{\partial S\left[\boldsymbol{x}+\lambda\boldsymbol{\epsilon}\right]}{\partial\lambda}\right)_{\boldsymbol{x},\boldsymbol{\epsilon}}\Bigg|_{\lambda=0}. \tag{26}$$

This is easiest to understand with an example. Consider the action

$$S\left[\boldsymbol{x}\right] = \int \mathrm{d}t' \sum_{m=1}^{N} \frac{1}{2} m\dot{x}_{m,t'}^2 - V\left(x_{m,t'}\right) \tag{27}$$

where I have used alternative dummy integration and summation variables t' and m; $\dot{x} = \mathrm{d}x/\mathrm{d}t'$ here. Now we have

$$\frac{\delta S\left[\boldsymbol{x}\right]}{\delta x_{n,t}} = \frac{\delta}{\delta x_{n,t}} \int \mathrm{d}t' \sum_{m=1}^{N} \frac{1}{2} m\dot{x}_{m,t'}^2 - V\left(x_{m,t'}\right) \tag{28}$$

$$= \int \mathrm{d}t' \sum_{m=1}^{N} \frac{1}{2} m \frac{\delta \dot{x}_{m,t'}^2}{\delta x_{n,t}} - \frac{\delta V\left(x_{m,t'}\right)}{\delta x_{n,t}}. \tag{29}$$

The chain rule gives

$$\frac{\delta S\left[\boldsymbol{x}\right]}{\delta x_{n,t}} = \int \mathrm{d}t' \sum_{m=1}^{N} m\dot{x}_{m,t'} \frac{\delta \dot{x}_{m,t'}}{\delta x_{n,t}} - \frac{\mathrm{d}V\left(x_{m,t'}\right)}{\mathrm{d}x_{m,t'}} \frac{\delta x_{m,t'}}{\delta x_{n,t}} \tag{30}$$

and integration by parts on the first term, using the boundary conditions, gives

$$\frac{\delta S\left[\boldsymbol{x}\right]}{\delta x_{n,t}} = \int \mathrm{d}t' \sum_{m=1}^{N} -m\ddot{x}_{m,t'} \frac{\delta x_{m,t'}}{\delta x_{n,t}} - \frac{\mathrm{d}V\left(x_{m,t'}\right)}{\mathrm{d}x_{m,t'}} \frac{\delta x_{m,t'}}{\delta x_{n,t}} \tag{31}$$

$$= -\int \mathrm{d}t' \sum_{m=1}^{N} \left(m\ddot{x}_{m,t'} + \frac{\mathrm{d}V\left(x_{m,t'}\right)}{\mathrm{d}x_{m,t'}}\right) \frac{\delta x_{m,t'}}{\delta x_{n,t}}. \tag{32}$$

Finally, we need the relation

$$\frac{\delta x_{m,t'}}{\delta x_{n,t}} = \delta_{mn}\delta\left(t-t'\right) \tag{33}$$

to eliminate the sum and integral, giving

$$\frac{\delta S\left[\boldsymbol{x}\right]}{\delta x_{n,t}} = -\int \mathrm{d}t' \sum_{m=1}^{N} \left(m\ddot{x}_{m,t'} + \frac{\mathrm{d}V\left(x_{m,t'}\right)}{\mathrm{d}x_{m,t'}}\right) \delta_{mn}\delta\left(t-t'\right) \tag{34}$$

$$= -\left(m\ddot{x}_{n,t} + \frac{\mathrm{d}V\left(x_{n,t}\right)}{\mathrm{d}x_{n,t}}\right) \tag{35}$$

where the dots now mean derivatives with respect to t. Setting the result equal to zero, according to the principal of least action, again gives

$$m\ddot{x}_{n,t} = -\frac{\mathrm{d}V\left(x_{n,t}\right)}{\mathrm{d}x_{n,t}} \tag{36}$$

as we found in Eq 25.

In general it is typically simpler to write functional derivatives out in full, as in the previous section, but the notation is convenient.

Note that the action is not directly a function of $x_{n,t}$, as n and t are summed/integrated over. Nevertheless, we take a functional derivative with respect to $x_{n,t}$. The effect is to eliminate the sum and integral, which promotes the respective n and t variables from dummy variables to free variables.

1.3 Hamiltonian Classical Mechanics

The Hamiltonian density is defined to be the Legendre tranform of the Lagrangian density:

$$\mathcal{H}\left(x_n, p_n\right) \triangleq p_n \dot{q}_n - \mathcal{L}\left(x_n, \dot{x}_n\right) \tag{37}$$

where the Hamiltonian is:

$$H = \sum_{n=1}^{N} \mathcal{H}\left(x_n, p_n\right). \tag{38}$$

In classical mechanics this is simply a function. The 'momentum conjugate to $\dot{x}_n$' is defined to be

$$p_{n,t} \triangleq \frac{\partial \mathcal{L}}{\partial \dot{x}_{n,t}} \tag{39}$$

where again I will typically neglect to write the time label explicitly. In Hamiltonian mechanics, the time evolution (dynamics) are found using

Hamilton's equations of motion:

$$\dot{x}_n = \left(\frac{\partial \mathcal{H}}{\partial p_n}\right)_{x_n}; \qquad \dot{p}_n = -\left(\frac{\partial \mathcal{H}}{\partial x_n}\right)_{p_n}. \tag{40}$$

1.4 Relativistic conventions

I will use the the $(+---)$ signature for the Minkowski metric:

$$\eta^{\mu\nu} = \eta_{\mu\nu} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}. \tag{41}$$

1.4.1 Covariant and Contravariant indices

Lorentz contra-variant 4-vectors are denoted with non-bold symbols with upper Greek indices

$$x^\mu = \begin{pmatrix} t \\ \boldsymbol{x} \end{pmatrix}^\mu; \quad p^\mu = \begin{pmatrix} E \\ \boldsymbol{p} \end{pmatrix}^\mu \tag{42}$$

and Lorentz co-variant 4-vectors[2] such as

$$\partial_\mu \triangleq \frac{\partial}{\partial x^\mu} \tag{43}$$

are denoted with lower Greek indices. Covectors can be created from contravectors by lowering the index with the Minkowski metric:

$$x_\mu = \eta_{\mu\nu} x^\nu \tag{44}$$

and vice versa:

$$x^\mu = \eta^{\mu\nu} x_\mu.$$

Under Lorentz transformations $\Lambda^\mu{}_\nu$ contravectors transform as

$$x^\mu \to (x')^\mu = \Lambda^\mu{}_\nu x^\nu \tag{45}$$

and covectors transform as

[2] A helpful mnemonic: "co is low".

$$x_\mu \to (x')_\mu = \Lambda_\mu{}^\nu x_\nu \tag{46}$$

where

$$\Lambda_\mu{}^\nu = \left(\Lambda^{-1}\right)^\mu{}_\nu . \tag{47}$$

E.g. a Lorentz boost in the x-direction is

$$\begin{pmatrix} t \\ x \\ y \\ z \end{pmatrix}^\mu \to \begin{pmatrix} t' \\ x' \\ y' \\ z' \end{pmatrix}^\mu = \begin{pmatrix} \gamma\left(t - \beta x\right) \\ \gamma\left(x - \beta t\right) \\ y \\ z \end{pmatrix}^\mu \tag{48}$$

$$= \begin{pmatrix} \gamma & -\beta\gamma & 0 & 0 \\ -\beta\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}^\mu{}_\nu \begin{pmatrix} t \\ x \\ y \\ z \end{pmatrix}^\nu \tag{49}$$

$$= \Lambda^\mu{}_\nu x^\nu . \tag{50}$$

Spacelike 3-vectors are typically denoted with bold symbols, although their individual components n (Latin indices) are non-bold:

$$[\boldsymbol{x}]_n = x_n . \tag{51}$$

Similarly, 1D spacelike vectors are non-bold, as in Eq 48. This will be important in Section 2 where we primarily consider a non-relativistic 1D chain.

1.4.2 Lorentz invariant quantities

The following integration measures are Lorentz invariant:

$$\int \mathrm{d}^4 x \triangleq \int \mathrm{d}t \int \mathrm{d}^3 \boldsymbol{x} \tag{52}$$

$$\int \mathrm{d}^4 p \triangleq \int \mathrm{d}E \int \mathrm{d}^3 \boldsymbol{p}. \tag{53}$$

To see this, note that a Lorentz transformation simply enacts a change of integration variables; hence, we need the Jacobian

$$\int \mathrm{d}^4 x \to \int \mathrm{d}^4 x' = \int \left|\det\left(\frac{\partial x'^\mu}{\partial x^\nu}\right)\right| \mathrm{d}^4 x = \int \left|\det\left(\Lambda^\mu{}_\nu\right)\right| \mathrm{d}^4 x \tag{54}$$

but

$$\left|\det\left(\Lambda^\mu{}_\nu\right)\right| = 1 \tag{55}$$

for any Lorentz transformation (they conserve spacetime volumes).
Einstein's equation

$$E^2 = m^2 + \boldsymbol{p}^2 \tag{56}$$

is of course also Lorentz invariant, and therefore so is any function of $E^2 - m^2 - \boldsymbol{p}^2$. In particular, the quantity

$$\delta\left(E^2 - m^2 - \boldsymbol{p}^2\right) \tag{57}$$

is Lorentz invariant, where we have defined

the *Dirac delta function*

$$\delta(x) \triangleq \frac{1}{2\pi}\int_{-\infty}^{\infty} \exp(ipx)\,\mathrm{d}p. \tag{58}$$

Heuristically,

$$\delta(x) \approx \begin{cases} 0, & x \neq y \\ \infty, & x = y \end{cases} \tag{59}$$

normalised such that

$$\int_{-\infty}^{\infty} \mathrm{d}x\delta(x) = 1. \tag{60}$$

As a result,

$$\int \mathrm{d}^4 p\delta\left(E^2 - m^2 - \boldsymbol{p}^2\right) \tag{61}$$

must also be Lorentz invariant. Carrying out the E integral in Eq 61 reveals that

$$\int \frac{\mathrm{d}^3\boldsymbol{p}}{2E_{\boldsymbol{p}}} \quad \text{is also a Lorentz invariant integration measure} \tag{62}$$

(the factor of 2 proves convenient later) where[3]

$$E_{\boldsymbol{p}} \triangleq \sqrt{m^2 + \boldsymbol{p}^2} \tag{63}$$
$$= \gamma m \tag{64}$$
$$= \omega_{\boldsymbol{p}}. \tag{65}$$

We can similarly define a

D-dimensional Dirac delta function:

$$\delta^D(\boldsymbol{x}) \triangleq (2\pi)^{-D}\int_{-\infty}^{\infty} \mathrm{d}^D\boldsymbol{p}\exp(i\boldsymbol{p}\cdot\boldsymbol{x}) \tag{66}$$

$$\approx \begin{cases} 0, & \boldsymbol{x} \neq \boldsymbol{y} \\ \infty, & \boldsymbol{x} = \boldsymbol{y} \end{cases} \tag{67}$$

$$\text{with } \int_{-\infty}^{\infty} \mathrm{d}^D\boldsymbol{x}\delta^D(\boldsymbol{x}) = 1. \tag{68}$$

Using this, we have that

$$\int \frac{\mathrm{d}^3\boldsymbol{p}}{2E_{\boldsymbol{p}}} 2E_{\boldsymbol{p}}\delta^3(\boldsymbol{p}) = 1 \tag{69}$$

and hence

$$2E_{\boldsymbol{p}}\delta^3(\boldsymbol{p}) \quad \text{is a Lorentz invariant 3D delta function.} \tag{70}$$

In non-relativistic quantum mechanics we know that two different momentum eigenstates are orthogonal:

$$\langle\boldsymbol{p}|\boldsymbol{q}\rangle = \delta^3(\boldsymbol{p} - \boldsymbol{q}) \tag{71}$$

hence it is natural to define

$$\text{4-momentum eigenstates} \quad |p\rangle \triangleq \sqrt{2E_{\boldsymbol{p}}}|\boldsymbol{p}\rangle \tag{72}$$

[3]Technically this also requires that choosing the positive energy branch is Lorentz invariant, which it is.

(where p is understood to be a 4-vector and $\boldsymbol{p}$ a 3-vector) so that

$$\langle p|q\rangle = \sqrt{2E_{\boldsymbol{p}}2E_{\boldsymbol{q}}}\langle \boldsymbol{p}|\boldsymbol{q}\rangle \tag{73}$$

$$= 2E_{\boldsymbol{p}}\delta^3\left(\boldsymbol{p}-\boldsymbol{q}\right) \tag{74}$$

which is the Lorentz invariant 3D delta function defined above. Note that while p^μ is contravariant, $|p\rangle$ does not carry an index: p is just a convenient label, and quantum states are neither covariant nor contravariant.

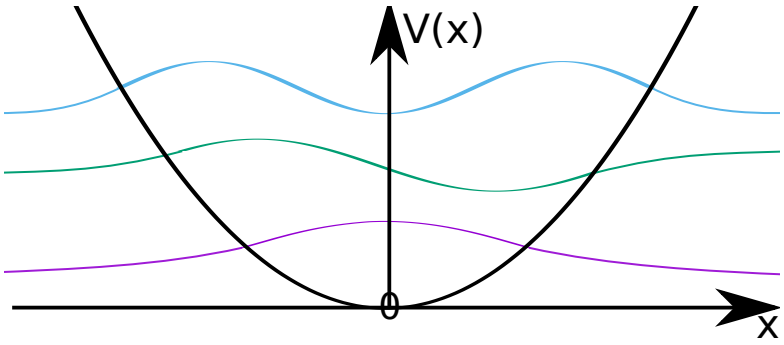


Figure 1: The quantum harmonic oscillator potential $V(x) = \frac{1}{2}m\omega^2 x^2$ and first few eigenstates.

2 Canonical Quantization of Fields

In this chapter we will see in full detail how to build a quantum field. We will focus on the simplest case of a **real, scalar, non-relativistic field**. As we will see, this will turn out to be a quantum theory of sound propagating through a crystal. This is the phonon field, whose particle excitations are called phonons.

References

- N. W. Ashcroft and N. D. Mermin, *Solid State Physics* (Harcourt 1976), Appendix L
- R. P. Feynman, *Statistical Mechanics* (Advanced Book Classics, 1972) Chapter 6
- A. Altland and B. Simons, Condensed Matter Field Theory (Cambridge University Press, 2010), Chapter 1
- T. Lancaster and S. J. Blundell, Quantum Field Theory for the Gifted Amateur (Oxford University Press, 2014), Section I
- S. M. Girvin and K. Yang, *Modern Condensed Matter Physics* (Cambridge University Press, 2019), Section 6.2

2.1 Recap: Quantum Harmonic Oscillators

Recall the Time Independent Schroedinger Equation (TISE) for the quantum harmonic oscillator:

$$\hat{H}|n\rangle = \left(\frac{\hat{p}^2}{2m} + \frac{1}{2}m\omega^2\hat{x}^2\right)|n\rangle = E_n|n\rangle. \tag{75}$$

The potential is shown in the position basis in Fig. 1.
An elegant way to find all eigenstates and eigenvalues is to use creation and annihilation operators (aka ladder operators, aka raising and lowering operators). The intuition developed in this simple QM problem forms the entire basis of QFT.

2.1.1 Creation and Annihilation Operators

The raising operator is

$$\boxed{\hat{a}^\dagger \triangleq \sqrt{\frac{m\omega}{2}}\left(\hat{x} - \frac{i}{m\omega}\hat{p}\right).} \tag{76}$$

It is non-Hermitian. Its hermitian conjugate is the lowering operator

$$\boxed{\hat{a} = \sqrt{\frac{m\omega}{2}}\left(\hat{x} + \frac{i}{m\omega}\hat{p}\right).} \tag{77}$$

This gives

$$\hat{a}^\dagger \hat{a} = \frac{m\omega}{2}\left(\hat{x}^2 + \frac{\hat{p}^2}{m^2\omega^2} + \frac{i}{m\omega}\left(\hat{x}\hat{p} - \hat{p}\hat{x}\right)\right). \tag{78}$$

The term in nested parentheses is just the commutator

$$[\hat{x}, \hat{p}] = i\hat{\mathbb{I}} \tag{79}$$

where $\hat{\mathbb{I}}$ is the identity operator. Hence,

$$\hat{a}^\dagger \hat{a} = \frac{m\omega}{2}\hat{x}^2 + \frac{\hat{p}^2}{2m\omega} - \frac{1}{2}\hat{\mathbb{I}}$$

and so

$$\omega\left(\hat{a}^\dagger \hat{a} + \frac{1}{2}\hat{\mathbb{I}}\right) = \hat{H}. \tag{80}$$

Therefore the TISE can be written

$$\boxed{\hat{H}|n\rangle = \omega\left(\hat{a}^\dagger \hat{a} + \frac{1}{2}\hat{\mathbb{I}}\right)|n\rangle = E_n|n\rangle.} \tag{81}$$

2.1.2 Commutation relations

The commutator of the creation and annihilation operators is

$$\left[\hat{a}, \hat{a}^\dagger\right] = \frac{m\omega}{2}\left[\hat{x} + \frac{i}{m\omega}\hat{p}, \hat{x} - \frac{i}{m\omega}\hat{p}\right] \tag{82}$$

$$= -i\left[\hat{x}, \hat{p}\right] \tag{83}$$

and so

$$\boxed{\left[\hat{a}, \hat{a}^\dagger\right] = \hat{\mathbb{I}}} \tag{84}$$

The commutator of the operators with the Hamiltonian is therefore:

$$\left[\hat{H}, \hat{a}^\dagger\right] = \left[\omega\left(\hat{a}^\dagger \hat{a} + \frac{1}{2}\hat{\mathbb{I}}\right), \hat{a}^\dagger\right] \tag{85}$$

$$= \omega\left(\hat{a}^\dagger \hat{a}\hat{a}^\dagger - \hat{a}^\dagger \hat{a}^\dagger \hat{a}\right) \tag{86}$$

$$= \omega\hat{a}^\dagger\left[\hat{a}, \hat{a}^\dagger\right] \tag{87}$$

giving

$$\boxed{\left[\hat{H}, \hat{a}^\dagger\right] = \omega\hat{a}^\dagger.} \tag{88}$$

Similarly,

$$\left[\hat{H}, \hat{a}\right] = -\omega\hat{a}. \tag{89}$$

2.1.3 Energy eigenstates and eigenvalues

Using Eq. 88 we can see the effect of $\hat{a}^\dagger$ on the TISE of Eq. 81:

$$\hat{H}|n\rangle = E_n|n\rangle \tag{90}$$

$$\downarrow$$

$$\hat{a}^\dagger\hat{H}|n\rangle = E_n\hat{a}^\dagger|n\rangle \tag{91}$$

$$\downarrow$$

$$\left(\hat{H}\hat{a}^\dagger - \left[\hat{H},\hat{a}^\dagger\right]\right)|n\rangle = E_n\hat{a}^\dagger|n\rangle \tag{92}$$

$$\downarrow$$

$$\left(\hat{H}\hat{a}^\dagger - \omega\hat{a}^\dagger\right)|n\rangle = E_n\hat{a}^\dagger|n\rangle \tag{93}$$

and the final result

$$\boxed{\hat{H}\left(\hat{a}^\dagger|n\rangle\right) = (E_n+\omega)\left(\hat{a}^\dagger|n\rangle\right).} \tag{94}$$

That is, if $|n\rangle$ is an eigenstate of the harmonic oscillator with eigenvalue E_n, then $\hat{a}^\dagger|n\rangle$ is an eigenstate with eigenvalue $E_n+\omega$. Repeating the process l times we find

$$\hat{H}\left(\left(\hat{a}^\dagger\right)^l|n\rangle\right) = (E_n+l\omega)\left(\left(\hat{a}^\dagger\right)^l|n\rangle\right). \tag{95}$$

This tells us that the energy levels are evenly spaced, and that

$$\left(\hat{a}^\dagger\right)^l|n\rangle \propto |n+l\rangle. \tag{96}$$

Similarly, we find

$$\hat{H}\left(\hat{a}|n\rangle\right) = (E_n-\omega)\left(\hat{a}|n\rangle\right). \tag{97}$$

The creation and annihilation operators move the state up or down the rungs of an energy ladder with evenly spaced rungs. While there exist an infinite number of rungs, the energies do not stretch down to negative energies. To see this, first note that

$$|\hat{a}|n\rangle|^2 \geq 0 \tag{98}$$

because the thing on the left, whatever it is, is the square modulus of something, and that is always ≥ 0. Expanding it out we have

$$\langle n|\hat{a}^\dagger\hat{a}|n\rangle \geq 0 \tag{99}$$

$$\langle n|\frac{1}{\omega}\hat{H} - \frac{1}{2}|n\rangle \geq 0 \tag{100}$$

$$\langle n|\frac{1}{\omega}E_n - \frac{1}{2}|n\rangle \geq 0 \tag{101}$$

$$E_n \geq \frac{\omega}{2}. \tag{102}$$

Therefore there is a lowest-energy state, a lowest rung to the ladder. This is the ground state which we denote $|0\rangle$. This is just a convenient label for a ket. It is <u>not</u> the number zero! You can think of it as shorthand for $|\psi_{n=0}\rangle$. The ground state has the property that

$$\boxed{\hat{a}|0\rangle = 0} \tag{103}$$

where the right hand side really *is* the number 0, so that any further action of lowering operators continues to return 0. To find the energy of the ground state we can work in the position basis using the definition of the lowering operator from Eq. 77:

$$\sqrt{\frac{m\omega}{2}}\left(\hat{x}+\frac{i}{m\omega}\hat{p}\right)|0\rangle=0 \tag{104}$$

$$\downarrow$$

$$\left(x+\frac{1}{m\omega}\frac{\mathrm{d}}{\mathrm{d}x}\right)\phi_0(x)=0. \tag{105}$$

This is a first order linear ODE which has the solution

$$\phi_0(x)\propto\exp\left(-\frac{m\omega}{2}x^2\right). \tag{106}$$

Inserting this back into the TISE gives

$$E_0=\frac{\omega}{2}. \tag{107}$$

This is the **ground state energy**, also called the **zero-point energy.**
Combining the results, we see that

$$\hat{H}|n\rangle=\omega\left(\hat{a}^\dagger\hat{a}+\frac{1}{2}\right)|n\rangle=\omega\left(n+\frac{1}{2}\right)|n\rangle$$

for integer n. Defining the 'number operator'

$$\hat{n}\triangleq\hat{a}^\dagger\hat{a} \tag{108}$$

we see that

$$\hat{n}|n\rangle=n|n\rangle. \tag{109}$$

2.1.4 Normalization

We have that

$$\langle n|\hat{n}|n\rangle=n \tag{110}$$

$$\downarrow$$

$$\langle n|\hat{a}^\dagger\hat{a}|n\rangle=n \tag{111}$$

$$\downarrow$$

$$\|\hat{a}|n\rangle\|^2=n \tag{112}$$

$$\downarrow$$

$$\|\hat{a}|n\rangle\|=\sqrt{n}. \tag{113}$$

Therefore, since

$$\hat{a}|n\rangle\propto|n-1\rangle$$

and

$$\langle n-1|n-1\rangle=1 \tag{114}$$

we have

$$\boxed{\hat{a}|n\rangle=\sqrt{n}|n-1\rangle.} \tag{115}$$

Similarly,

$$\left|\hat{a}^\dagger|n\rangle\right|^2 = \langle n|\hat{a}\hat{a}^\dagger|n\rangle = \langle n|\hat{n} + \left[\hat{a}, \hat{a}^\dagger\right]|n\rangle = n+1 \tag{116}$$

and so

$$\boxed{\hat{a}^\dagger|n\rangle = \sqrt{n+1}|n+1\rangle} \,. \tag{117}$$

By induction we can also see that

$$\boxed{\frac{1}{\sqrt{n!}}\left(\hat{a}^\dagger\right)^n|0\rangle = |n\rangle.} \tag{118}$$

2.2 Second quantization

- Equation 118 tells us we can create the n^{th} excited state of the harmonic oscillator by acting n raising operators on the ground state.
- Since all rungs of the ladder are evenly spaced with spacing ω, we can also interpret the n^{th} excited state as the presence of n identical particles each of energy ω.
- This is the origin of the name 'creation and annihilation operators': they create and annihilate particles in a second-quantised description.
- 'First quantization' is the realisation that classical particles exhibit wave-like properties in QM. We call the process of rewriting a problem in terms of creation and annihilation operators 'second quantization'. It is the realisation that classical waves gain particle-like properties in QM. The descriptions are equivalent.
- The fact that we can fit multiple particles into the same state identifies $\hat{a}^\dagger$ as the creation operator of a *boson*.

2.3 N Isolated Quantum Harmonic Oscillators

Now consider N independent QHOs evenly spaced with spacing a along a chain. Assume periodic boundary conditions, so that site $N+1$ is equivalent to site 1 (*i.e.* the chain is really a ring). The Hamiltonian is

$$\hat{H} = \sum_{n=1}^{N} \frac{\hat{p}_n^2}{2m} + \frac{1}{2}m\omega^2\hat{x}_n^2. \tag{119}$$

We can rewrite this using creation and annihilation operators, each with a site label:

$$\hat{H} = \sum_{n=1}^{N} \omega\left(\hat{a}_n^\dagger\hat{a}_n + \frac{1}{2}\hat{\mathbb{I}}\right). \tag{120}$$

Here, $\hat{a}_n^\dagger$ creates an excitation of energy ω at site n. We can also take advantage of periodicity (translational symmetry) to Fourier transform, using

$$\hat{a}_n^\dagger = \frac{1}{\sqrt{N}}\sum_{k=1}^{N} \exp\left(2\pi i\frac{k}{N}n\right)\hat{a}_k^\dagger \tag{121}$$

$$\hat{a}_k^\dagger = \frac{1}{\sqrt{N}}\sum_{n=1}^{N} \exp\left(-2\pi i\frac{k}{N}n\right)\hat{a}_n^\dagger \tag{122}$$

Figure 2: A classical chain of balls and springs.

to give

$$\hat{H} = \sum_n \omega \hat{a}_n^\dagger \hat{a}_n + \omega \frac{1}{2}\hat{\mathbb{I}}N \tag{123}$$

$$= \frac{1}{N}\sum_{n,k,q} \omega \exp\left(2\pi i \frac{k-q}{N} n\right) \hat{a}_k^\dagger \hat{a}_q + \omega \frac{1}{2}\hat{\mathbb{I}}N$$

$$= \sum_{k,q} \omega \delta_{kq} \hat{a}_k^\dagger \hat{a}_q + \omega \frac{1}{2}\hat{\mathbb{I}}N \tag{124}$$

$$= \sum_k \omega \hat{a}_k^\dagger \hat{a}_k + \omega \frac{1}{2}\hat{\mathbb{I}}N \tag{125}$$

where we used the useful expression

$$\frac{1}{N}\sum_{n=1}^{N} \exp\left(2\pi i \frac{n}{N}(k-q)\right) = \delta_{kq} \tag{126}$$

which follows since the sum is over equally distributed points around the unit circle. In fact we can also use this expression to rewrite as

$$\hat{H} = \sum_k^N \omega \left(\hat{a}_k^\dagger \hat{a}_k + \frac{1}{2}\hat{\mathbb{I}}\right). \tag{127}$$

In this case, $\hat{a}_k^\dagger$ creates an excitation of energy ω with wavevector $k_n = \pi k/N$, where $1 \leq k \leq N$.
The Hamiltonian in Eq 119 is overly simplified: the QHOs are uncoupled, and so it is hard to imagine how one might physically create an excitation across multiple oscillators. It is a remarkable result, which we will derive shortly, that when interactions are turned on between the oscillators, provided translational invariance is preserved, the effect is merely to add a wavevector dependence to the angular frequency $\omega \to \omega_k$. That is:

$$\hat{H}_{\text{interacting}} = \sum_k^N \omega_k \left(\hat{a}_k^\dagger \hat{a}_k + \frac{1}{2}\hat{\mathbb{I}}\right). \tag{128}$$

Before seeing how this works, let us first build intuition by looking at the classical problem.

2.4 N Coupled Classical Harmonic Oscillators

Consider a classical chain of balls connected by springs of stiffness K (Fig. 2) along an N-site chain. This is a chain of coupled classical harmonic oscillators. The balls oscillate about equilibrium positions an (where a is a lattice constant and $n \in [1, N]$). Let the position of ball n at time t be $u_{n,t}$, and define the displacement from the equilibrium position to be $x_{n,t} = an + u_{n,t}$. From now on we will only care about the small displacements $u_{n,t}$.

2.4.1 Recap: forces on balls and springs

Wobbling one of the balls will lead to a vibration propagating through the chain. To find the dispersion relation of this propagating mode, consider the force on the n^{th} ball (using Newton's 2nd law):

$$m\ddot{x}_{n,t} = K\left(x_{n+1,t} - x_{n,t}\right) - K\left(x_{n,t} - x_{n-1,t}\right). \tag{129}$$

We need the normal modes of oscillation. To find them, first take advantage of the translational symmetry by applying a Fourier transform:

$$x_{n,t} = \frac{1}{\sqrt{N}} \sum_{k=1}^{N} \int \mathrm{d}\omega \exp\left(2\pi i \left(\frac{kn}{N} - \omega t\right)\right) Q_{k,\omega} \tag{130}$$

(note that the sum is over integer ka, so that k now has units of inverse length) to give

$$\frac{m}{\sqrt{N}} \sum_{k=1}^{N} \int \mathrm{d}\omega \exp\left(2\pi i \left(kn/N - \omega t\right)\right) \left(-\omega^2\right) Q_{k,\omega} = \frac{K}{\sqrt{N}} \sum_{k=1}^{N} \int \mathrm{d}\omega \exp\left(2\pi i \left(kn/N - \omega t\right)\right) \left(Q_{k,\omega} \left(\exp\left(2\pi i k/N\right) - 1\right) - (1 - \right. \tag{131}$$

We can inverse Fourier transform to find a separate solution for each pair of k and ω:

$$-m\omega_k^2 Q_{k,\omega} = 2K \left(\cos\left(2\pi k/N\right) - 1\right) Q_{k,\omega} \tag{132}$$

where the notation ω_k emphasises that this is a specific solution $\omega\left(k\right)$. Hence

$$\omega_k = \sqrt{\frac{2K}{m}\left(1 - \cos\left(2\pi k/N\right)\right)} \tag{133}$$

$$= 2\sqrt{\frac{K}{m}} \left|\sin\left(\frac{\pi k}{N}\right)\right| \tag{134}$$

(noting that only the positive root is physical). Around $k \approx 0$ (long wavelengths) the mode is linearly dispersing:

$$\omega_k \approx \sqrt{\frac{K}{m}} \left|k\right| \tag{135}$$

which defines the speed of sound $\sqrt{\frac{K}{m}}$ along the chain.

2.4.2 The classical Lagrangian

Thanks to work carried out in Bristol by Aharonov and Bohm, we know that in quantum mechanics we must deal with potentials rather than forces. To set up the quantum problem we must rephrase the classical problem construct the Lagrangian, and from there the Hamiltonian, of the classical system. We will then quantize this in Section 2.5.
The classical balls and springs model is governed by the Lagrangian density

$$\mathcal{L} = \frac{1}{2} m \dot{x}_n^2 - \frac{1}{2} m\omega^2 \left(x_{n+1} - x_n\right)^2 - \frac{1}{2} m\omega^2 \left(x_n - x_{n-1}\right)^2 . \tag{136}$$

To check this is equivalent to Eq. 129, we can derive the classical equations of motion using the Euler Lagrange equations:

$$\left(\frac{\partial \mathcal{L}}{\partial x_n}\right)_{\dot{x}_n} - \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial \mathcal{L}}{\partial \dot{x}_n}\right)_{x_n} = 0 \tag{137}$$

along with the identities

$$\left(\frac{\partial x_m}{\partial x_n}\right)_{\dot{x}_m} = \left(\frac{\partial \dot{x}_m}{\partial \dot{x}_n}\right)_{x_m} = \delta_{nm} \tag{138}$$

where we have introduced the

Kronecker delta:

$$\delta_{nm} \triangleq \begin{cases} 0, & n \neq m \\ 1, & n = m. \end{cases} \tag{139}$$

We find

$$\left(\frac{\partial \mathcal{L}}{\partial x_n}\right)_{\dot{x}_n} = \sum_{m=1}^{N} -m\omega^2 \left(x_{m+1} - x_m\right)\left(\delta_{m+1,n} - \delta_{m,n}\right) - m\omega^2 \left(x_m - x_{m-1}\right)\left(\delta_{m,n} - \delta_{m-1,n}\right) \tag{140}$$

$$= -2m\omega^2 \left(2x_n - x_{n-1} - x_{n+1}\right) \tag{141}$$

and

$$\left(\frac{\partial \mathcal{L}}{\partial \dot{x}_n}\right)_{x_n} = m\dot{x}_n \tag{142}$$

so that the Euler Lagrange equations read

$$-2m\omega^2 \left(2x_n - x_{n-1} - x_{n+1}\right) - m\ddot{x}_n = 0 \tag{143}$$

$$\downarrow \tag{144}$$

$$m\ddot{x}_n = 2m\omega^2 \left(x_{n+1} - x_n\right) - 2m\omega^2 \left(x_n - x_{n-1}\right) \tag{145}$$

which, as expected, is Eq. 129 (with force constant $K = 2m\omega^2$).

2.4.3 The Classical Hamiltonian

To find the Hamiltonian, we need the momentum canonically conjugate to the position. This is defined to be:

$$p_n \triangleq \left(\frac{\partial \mathcal{L}}{\partial \dot{x}_n}\right)_{x_n} \tag{146}$$

which in this case gives

$$p_n = m\dot{x}_n. \tag{147}$$

The Hamiltonian is then the Legendre transform of the Lagrangian:

$$H \triangleq \sum_n p_n \cdot \dot{x}_n - L. \tag{148}$$

In this case we find

$$H = \sum_n \frac{1}{2m} p_n^2 + \frac{1}{2} m\omega^2 \left(x_{n+1} - x_n\right)^2 + \frac{1}{2} m\omega^2 \left(x_n - x_{n-1}\right)^2. \tag{149}$$

This is the classical Hamiltonian for the balls and springs model. It is a constant of motion corresponding to the total energy of the system.

2.5 N Coupled Quantum Harmonic Oscillators

Now let's combine these ideas, with a chain of coupled quantum harmonic oscillators. Our aim will be to confirm the intuition that we can decompose the problem into one of N independent QHOs, one for each wavevector k.

To create the quantum problem we can 'canonically quantize' the classical Hamiltonian in Eq. 149. This simply means promoting observables to operators:

$$\hat{H} = \sum_{n=1}^{N} \frac{\hat{p_n}^2}{2m} + \frac{1}{2} m\omega^2 \sum_n \left(\hat{x}_{n+1} - \hat{x}_n\right)^2 + \left(\hat{x}_n - \hat{x}_{n-1}\right)^2 \tag{150}$$

which obey the 'canonical commutation relations':

$$[\hat{x}_n, \hat{p}_m] = i\hat{\mathbb{I}}\delta_{nm} \tag{151}$$
$$[\hat{x}_n, \hat{x}_m] = 0 \tag{152}$$
$$[\hat{p}_n, \hat{p}_m] = 0. \tag{153}$$

Note that operators on different sites always commute. We can again take advantage of translational invariance to define

$$\hat{x}_n = \frac{1}{\sqrt{N}} \sum_{k=1}^{N} \exp\left(2\pi i k n/N\right) \hat{Q}_k \tag{154}$$
$$\hat{p}_n = \frac{1}{\sqrt{N}} \sum_{k=1}^{N} \exp\left(-2\pi i k n/N\right) \hat{\Pi}_k \tag{155}$$

and

$$\hat{Q}_k = \frac{1}{\sqrt{N}} \sum_{n=1}^{N} \exp\left(-2\pi i k n/N\right) \hat{x}_n \tag{156}$$
$$\hat{\Pi}_k = \frac{1}{\sqrt{N}} \sum_{n=1}^{N} \exp\left(2\pi i k n/N\right) \hat{p}_n \tag{157}$$

where I have reverted to summing over integer k (rather than integer ka). Note that these tranformed co-ordinates also obey canonical commutation relations:

$$\left[\hat{Q}_k, \hat{\Pi}_q\right] = \frac{1}{N} \sum_{n,m} \exp\left(2\pi i \left(qm - kn\right)/N\right) \left[\hat{x}_n, \hat{p}_m\right] \tag{158}$$
$$= \frac{1}{N} \sum_{n,m} \exp\left(2\pi i \left(qm - kn\right)/N\right) i\hbar\hat{\mathbb{I}}\delta_{nm} \tag{159}$$
$$= \frac{1}{N} \sum_{n} \exp\left(2\pi i \left(q - k\right) n/N\right) i\hbar\hat{\mathbb{I}} \tag{160}$$
$$= i\hat{\mathbb{I}}\delta_{kq} \tag{161}$$

using Eq 126 in the last line. Similarly,

$$\left[\hat{Q}_k, \hat{Q}_q\right] = \left[\hat{\Pi}_k, \hat{\Pi}_q\right] = 0. \tag{162}$$

However, $\hat{Q}_k$ and $\hat{\Pi}_k$ are no longer Hermitian. Rather,

$$\hat{Q}_k^\dagger = \hat{Q}_{-k}. \tag{163}$$

In these new variables the Hamiltonian reads

$$\hat{H} = \frac{1}{2mN} \sum_{n,k,q} \exp\left(-2\pi i \left(k + q\right) n/N\right) \hat{\Pi}_k \hat{\Pi}_q \tag{164}$$
$$+ \frac{m\omega^2}{2N} \sum_{n,k,q} \hat{Q}_k \hat{Q}_q \exp\left(2\pi i \left(k + q\right) n/N\right) \left(\exp\left(2\pi i \left(k + q\right)/N\right) + 2 + \exp\left(-2\pi i \left(k + q\right)/N\right) - 2\exp\left(2\pi i k/N\right) - 2\exp\left(-2\right.\right. \tag{165}$$

where I have expanded the potential terms. Hence

$$\hat{H} = \frac{1}{2m}\sum_k \hat{\Pi}_{-k}\hat{\Pi}_k + 4m\omega^2 \sum_k \sin^2\left(\frac{\pi k}{N}\right)\hat{Q}_{-k}\hat{Q}_k \tag{166}$$

or

$$\hat{H} = \sum_k \frac{1}{2m}\hat{\Pi}_{-k}\hat{\Pi}_k + \frac{1}{2}m\omega_k^2\hat{Q}_{-k}\hat{Q}_k \tag{167}$$

$$= \frac{1}{2m}\hat{\Pi}_k^\dagger\hat{\Pi}_k + \frac{1}{2}m\omega_k^2\hat{Q}_k^\dagger\hat{Q}_k \tag{168}$$

where

$$\omega_k \triangleq \sqrt{8}\omega\left|\sin\left(\frac{\pi k}{N}\right)\right|. \tag{169}$$

Unfortunately Eq. 167 makes it clear that states at k and $-k$ are coupled, which is not quite what we want. If we can decouple the modes, we can find the eigenstates and eigenvalues using creation and annihilation operators as for the single harmonic oscillator. A clever definition of creation operator allows us to do this while simultaneously decoupling k from $-k$ (the two steps can also be done separately, with more algebra). We define a different creation operator for each mode k:

$$\boxed{\hat{a}_k^\dagger \triangleq \sqrt{\frac{m\omega_k}{2}}\left(\hat{Q}_{-k} - \frac{i}{m\omega_k}\hat{\Pi}_k\right)} \tag{170}$$

giving

$$\hat{a}_k = \sqrt{\frac{m\omega_k}{2}}\left(\hat{Q}_{-k}^\dagger + \frac{i}{m\omega_k}\hat{\Pi}_k^\dagger\right) \tag{171}$$

$$= \sqrt{\frac{m\omega_k}{2}}\left(\hat{Q}_k + \frac{i}{m\omega_k}\hat{\Pi}_{-k}\right). \tag{172}$$

Note that $\hat{a}_k^\dagger \neq \hat{a}_{-k}$. These obey the commutation relations:

$$\left[\hat{a}_k, \hat{a}_q^\dagger\right] = \frac{m\omega_k}{2}\left[\hat{Q}_k + \frac{i}{m\omega_k}\hat{\Pi}_{-k}, \hat{Q}_{-q} - \frac{i}{m\omega_k}\hat{\Pi}_q\right] \tag{173}$$

$$= \frac{m\omega_k}{2}\left(\left[\hat{Q}_k, -\frac{i}{m\omega_k}\hat{\Pi}_q\right] + \left[\frac{i}{m\omega_k}\hat{\Pi}_{-k}, \hat{Q}_{-q}\right]\right) \tag{174}$$

$$= \frac{i}{2}\left(\left[\hat{\Pi}_{-k}, \hat{Q}_{-q}\right] - \left[\hat{Q}_k, \hat{\Pi}_q\right]\right) \tag{175}$$

$$= \delta_{kq}\hat{\mathbb{I}} \tag{176}$$

and

$$[\hat{a}_k, \hat{a}_q] = \left[\hat{a}_k^\dagger, \hat{a}_q^\dagger\right] = 0. \tag{177}$$

We can now re-arrange to get

$$\hat{Q}_k = \sqrt{\frac{1}{2m\omega_k}}\left(\hat{a}_k + \hat{a}_{-k}^\dagger\right) \tag{178}$$

$$\hat{\Pi}_k = i\sqrt{\frac{m\omega_k}{2}}\left(\hat{a}_k^\dagger - \hat{a}_{-k}\right) \tag{179}$$

and so

$$\hat{H} = \sum_{k=1}^{N} \frac{1}{2m}\hat{\Pi}_k^\dagger \hat{\Pi}_k + \frac{1}{2} m\omega_k^2 \hat{Q}_k^\dagger \hat{Q}_k \tag{180}$$

$$= \sum_{k=1}^{N} \frac{\omega_k}{4} \left\{ \left(\hat{a}_k - \hat{a}_{-k}^\dagger\right)\left(\hat{a}_k^\dagger - \hat{a}_{-k}\right) + \left(\hat{a}_k^\dagger + \hat{a}_{-k}\right)\left(\hat{a}_k + \hat{a}_{-k}^\dagger\right) \right\} \tag{181}$$

$$= \sum_{k=1}^{N} \frac{\omega_k}{4} \left(\hat{a}_k \hat{a}_k^\dagger + \hat{a}_{-k}^\dagger \hat{a}_{-k} + \hat{a}_k^\dagger \hat{a}_k + \hat{a}_{-k}\hat{a}_{-k}^\dagger \right) \tag{182}$$

$$= \sum_{k=1}^{N} \omega_k \left(\frac{1}{2}\hat{a}_k^\dagger \hat{a}_k + \frac{1}{2}\hat{a}_{-k}^\dagger \hat{a}_{-k} + \frac{1}{2}\hat{\mathbb{I}} \right) \tag{183}$$

where I have used the commutators. The modes are now nicely decoupled (k and $-k$ terms appear separately). As a result, noting that k is periodic (as is n), we can label states in such a manner that

$$\hat{H} = \sum_{k=1}^{N} \omega_k \left(\hat{a}_k^\dagger \hat{a}_k + \frac{1}{2}\hat{\mathbb{I}} \right). \tag{184}$$

Hence, using the Fourier transform we have written the coupled QHOs in terms of independent creation and annihilation operators.
The number operator

$$\hat{n}_k \triangleq \hat{a}_k^\dagger \hat{a}_k \tag{185}$$

counts the number of particles in mode k, i.e. with wavevector $\pi k/N$.

2.5.1 Excitations of a Quantum Chain: Fock Space

Each operator $\hat{a}_k^\dagger$ works exactly like a creation operator acts on a single QHO. Acting it on the ground state creates an excitation with a well defined wavevector. The ground state, or vacuum state, $|\Omega\rangle$ is a product of N independent modes:

$$|\Omega\rangle = |0_{k=1}\rangle \otimes |0_{k=2}\rangle \otimes |0_{k=3}\rangle \otimes |0_{k=4}\rangle \otimes \dots \tag{186}$$

which can be written compactly as

$$|\Omega\rangle = |0000\dots\rangle. \tag{187}$$

Acting a creation operator with wavevector k raises mode k, at a cost of ω. *E.g.*

$$\hat{a}_{k=3}^\dagger |\Omega\rangle = |00100\dots\rangle \tag{188}$$

or

$$\left(\hat{a}_{k=3}^\dagger\right)^3 \left(\hat{a}_{k=2}^\dagger\right)^4 \left(\hat{a}_{k=1}^\dagger\right)^2 |\Omega\rangle = |24300\dots\rangle \tag{189}$$

and so on. There's no problem with exciting any given mode multiple times, just as there's no problem with exciting a single QHO to its n^{th} energy level. Note that $|24300\dots\rangle$ is just shorthand for a tensor product of N separate kets:

$$|24300\dots\rangle \triangleq |2_{k=1}\rangle \otimes |4_{k=2}\rangle |3_{k=3}\rangle |0_{k=4}\rangle \dots \tag{190}$$

Hence, it is not a ket living in a Hilbert space, as in single-particle QM. Rather, we say that this object lives in a *Fock Space*. This is the space of all Hilbert spaces, each of which has a fixed particle number. You can equally well create an excitation at a specific position an by acting the creation operator

$$\hat{a}_n^\dagger = \frac{1}{\sqrt{N}} \sum_{k=1}^{N} \exp\left(2\pi i k n/N\right) \hat{a}_k^\dagger \tag{191}$$

Figure 3: Coarse graining positions into a continuum field.

on the vacuum state:

$$\hat{a}_n^\dagger|\Omega\rangle = \frac{1}{\sqrt{N}}\sum_{k=1}^{N}\exp\left(2\pi ikn/N\right)\hat{a}_k^\dagger|\Omega\rangle \tag{192}$$

$$= \frac{1}{\sqrt{N}}\exp\left(2\pi in/N\right)\hat{a}_{k=1}^\dagger|000\ldots\rangle + \exp\left(4\pi in/N\right)\hat{a}_{k=2}^\dagger|000\ldots\rangle + \ldots \tag{193}$$

$$= \frac{1}{\sqrt{N}}\exp\left(2\pi in/N\right)|100\ldots\rangle + \frac{1}{\sqrt{N}}\exp\left(4\pi in/N\right)|010\ldots\rangle + \ldots \tag{194}$$

You can think by analogy of plucking a guitar string at a given position along its length. The effect is to excite all the harmonics of the string with different amplitudes.

2.6 Classical Field Theory

So far we have worked with discrete positions n. A field is a continuum, so to proceed we must smooth out our description. Our motivation is that we are typically only interested in lengthscales / wavelengths significantly larger than a lattice spacing, and so we can 'coarse grain' or smooth over the lattice details.

2.6.1 Classical Lagrangian Field Theory

In the classical balls and springs model we can make a conceptual leap by approximating our discrete positions by a continuum field. This amounts to making the change

$$x_{n,t} \to \varphi_{\boldsymbol{x},t}. \tag{195}$$

have introduced bold $\boldsymbol{x}$ to emphasise that this works in any spatial dimension, and because we will shortly need to differentiate between 3-vectors $\boldsymbol{x}$ and 4-vectors $x^\mu = (t, \boldsymbol{x})^\mu$.
Formula 195 is the defining expression of field theory. While conceptually simple, the consequences are profound. Rather than individual balls oscillating about discrete sites n with displacements $x_{n,t}$, we instead have a continuous field defined at all positions $\boldsymbol{x}$ and times t. You can think of it as a flexible membrane, like the surface of a drum. See Fig 3. In crystals the result leads to simplified mathematics. In the standard model, however, there *are no individual sites*: only the smooth field exists, and it does so across all of spacetime. It is worth taking the time to think about Eq. 195 conceptually before proceeding.

The following might help (or it might do the opposite!). In classical mechanics it is common to denote generalised co-ordinates as $q_{n,t}$ rather than $x_{n,t}$. In this alternative notation, when we coarse grain and smooth out, we simply switch $q_{n,t} \to q_{\boldsymbol{x},t}$, interpolating the discrete variable n with continuous positions $\boldsymbol{x}$. Then we just relabel $q_{\boldsymbol{x},t} \to \varphi_{\boldsymbol{x},t}$ as a matter of convention.

The Lagrangian density coarse-grains as follows:

$$\mathcal{L} = \frac{1}{2}m\dot{x}_n^2 - \frac{1}{2}m\omega^2\left(x_{n+1} - x_n\right)^2 - \frac{1}{2}m\omega^2\left(x_n - x_{n-1}\right)^2 \tag{196}$$

$$\downarrow \text{coarse-grain} \tag{197}$$

$$\mathcal{L} = \frac{1}{2}m\dot{\varphi}_{\boldsymbol{x}}^2 - m\omega^2\left(\nabla\varphi_{\boldsymbol{x}}\right)^2. \tag{198}$$

Let us write things in a general spatial dimension D, as it is no more difficult that using $D = 1$ and it will allow us to use bold $\boldsymbol{x}$ (which will be helpful when we go relativistic). As before, neglect the t label since everything is evaluated at the same time. The Lagrangian is

$$L = \int \mathrm{d}^D\boldsymbol{x}\mathcal{L}\left(\varphi_{\boldsymbol{x}}, \dot{\varphi}_{\boldsymbol{x}}\right) \tag{199}$$

and action

$$S = \int \mathrm{d}t L = \int \mathrm{d}^{D+1}\boldsymbol{x}\mathcal{L}. \tag{200}$$

This notation already emphasises that the action is a Lorentz invariant scalar. We can find the Euler Lagrange equations using the principle of least action:

$$S\left[\varphi + \lambda\epsilon\right] = \int \mathrm{d}t \int \mathrm{d}^D\boldsymbol{x} \left\{ \frac{1}{2} m \left(\dot{\varphi_{\boldsymbol{x}}} + \lambda\dot{\epsilon_{\boldsymbol{x}}}\right)^2 - m\omega^2 \left(\nabla\varphi_{\boldsymbol{x}} + \lambda\nabla\epsilon_{\boldsymbol{x}}\right)^2 \right\} \tag{201}$$

$$\Downarrow \tag{202}$$

$$\left(\frac{\partial S\left[\varphi + \lambda\epsilon\right]}{\partial\lambda}\right)_{\varphi,\epsilon}\Bigg|_{\lambda=0} = \int \mathrm{d}t \int \mathrm{d}^D\boldsymbol{x} \left\{ m\dot{\varphi_{\boldsymbol{x}}}\dot{\epsilon_{\boldsymbol{x}}} - 2m\omega^2\nabla\varphi_{\boldsymbol{x}}\nabla\epsilon_{\boldsymbol{x}} \right\} \tag{203}$$

$$= \int \mathrm{d}t \int \mathrm{d}^D\boldsymbol{x}\epsilon_{\boldsymbol{x}} \left\{ -m\ddot{\varphi}_{\boldsymbol{x}} + 2m\omega^2\nabla^2\varphi_{\boldsymbol{x}} \right\} \tag{204}$$

using integration by parts with respect to time in the first term, and space in the second. Setting this equal to zero for all ϵ requires that

$$\ddot{\varphi}_{\boldsymbol{x}} = c^2\nabla^2\varphi_{\boldsymbol{x}} \tag{205}$$

where $c = \sqrt{2}\omega$. This is the classical wave equation. As hoped, we still have our classical sound mode propagating in our coarse-grained theory.

2.6.2 Classical Hamiltonian Field Theory

As before, we need the momentum conjugate to the field. This is given by

$$\pi_{\boldsymbol{x},t} \triangleq \left(\frac{\partial\mathcal{L}}{\partial\dot{\varphi}_{\boldsymbol{x},t}}\right)_{\varphi_{\boldsymbol{x},t}} \tag{206}$$

where we will continue to neglect the t label. In our case we find

$$\pi_{\boldsymbol{x}} = m\dot{\varphi}_{\boldsymbol{x}}. \tag{207}$$

The Hamiltonian is again the Legendre transform of the Lagrangian:

$$H \triangleq \int \mathrm{d}^D\boldsymbol{x}\pi_{\boldsymbol{x}}\dot{\varphi}_{\boldsymbol{x}} - L \tag{208}$$

and in our case this gives

$$H = \int \mathrm{d}^D\boldsymbol{x}\frac{1}{2m}\pi_{\boldsymbol{x}}^2 + m\omega^2 \int \mathrm{d}^D\boldsymbol{x} \left(\nabla\varphi_{\boldsymbol{x}}\right)^2. \tag{209}$$

2.7 Quantum Field Theory of the Harmonic Chain

2.7.1 Canonical Field Quantization

Finally we are ready to create our first quantum field. To do so, we simply promote our classical fields to non-commuting quantum operators:

$$\varphi_{\boldsymbol{x}} \to \hat{\varphi}_{\boldsymbol{x}} \tag{210}$$

$$\pi_{\boldsymbol{x}} \to \hat{\pi}_{\boldsymbol{x}}. \tag{211}$$

We are now explicitly justified in neglecting the t label since, in the Schroedinger picture, operators are time-independent. The quantum fields obey the canonical commutation relations

$$[\hat{\varphi}_{\boldsymbol{x}}, \hat{\varphi}_{\boldsymbol{y}}] = [\hat{\pi}_{\boldsymbol{x}}, \hat{\pi}_{\boldsymbol{y}}] = 0 \tag{212}$$

and

$$[\hat{\varphi}_{\boldsymbol{x}}, \hat{\pi}_{\boldsymbol{y}}] = i\delta^D(\boldsymbol{x}-\boldsymbol{y})\,\hat{\mathbb{I}}. \tag{213}$$

Note that in this description, unlike in single-particle quantum mechanics, positions are not themselves operators. Rather, positions should be thought of as like the fixed discrete lattice sites n, while the field is like x_n, the displacement of the ball from that site.
Our potential term has translational symmetry (as it only cares about the gradient of the field). We can again take advantage of this by working in the Fourier basis:

$$\hat{\varphi}_{\boldsymbol{x}} = \int \mathrm{d}^D\boldsymbol{k} \exp(2\pi i \boldsymbol{k}\cdot\boldsymbol{x})\,\hat{\Phi}_{\boldsymbol{k}} \tag{214}$$

$$\hat{\pi}_{\boldsymbol{x}} = \int \mathrm{d}^D\boldsymbol{k} \exp(-2\pi i \boldsymbol{k}\cdot\boldsymbol{x})\,\hat{\Pi}_{\boldsymbol{k}} \tag{215}$$

and the inverse transforms

$$\hat{\Phi}_{\boldsymbol{k}} = \int \mathrm{d}^D\boldsymbol{x} \exp(-2\pi i \boldsymbol{k}\cdot\boldsymbol{x})\,\hat{\varphi}_{\boldsymbol{x}} \tag{216}$$

$$\hat{\Pi}_{\boldsymbol{k}} = \int \mathrm{d}^D\boldsymbol{x} \exp(2\pi i \boldsymbol{k}\cdot\boldsymbol{x})\,\hat{\pi}_{\boldsymbol{x}}. \tag{217}$$

The new fields obey the canonical commutation relations

$$\left[\hat{\Phi}_{\boldsymbol{k}}, \hat{\Phi}_{\boldsymbol{q}}\right] = \left[\hat{\Pi}_{\boldsymbol{k}}, \hat{\Pi}_{\boldsymbol{q}}\right] = 0$$

and

$$\left[\hat{\Phi}_{\boldsymbol{k}}, \hat{\Pi}_{\boldsymbol{q}}\right] = \int \mathrm{d}^D\boldsymbol{x} \int \mathrm{d}^D\boldsymbol{y} \exp(2\pi i(\boldsymbol{q}\cdot\boldsymbol{y} - \boldsymbol{k}\cdot\boldsymbol{x}))\,[\hat{\varphi}_{\boldsymbol{x}}, \hat{\pi}_{\boldsymbol{y}}] \tag{218}$$

$$= \int \mathrm{d}^D\boldsymbol{x} \int \mathrm{d}^D\boldsymbol{y} \exp(2\pi i(\boldsymbol{q}\cdot\boldsymbol{y} - \boldsymbol{k}\cdot\boldsymbol{x}))\, i\delta^D(\boldsymbol{x}-\boldsymbol{y})\,\hat{\mathbb{I}} \tag{219}$$

$$= \int \mathrm{d}^D\boldsymbol{x} \exp(2\pi i(\boldsymbol{q}-\boldsymbol{k})\cdot\boldsymbol{x})\, i\hat{\mathbb{I}} \tag{220}$$

$$= i\delta^D(\boldsymbol{k}-\boldsymbol{q})\,\hat{\mathbb{I}} \tag{221}$$

using the useful relation

$$\boxed{\int \mathrm{d}^D\boldsymbol{x} \exp(2\pi i(\boldsymbol{q}-\boldsymbol{k})\cdot\boldsymbol{x}) = \delta^D(\boldsymbol{k}-\boldsymbol{q}).} \tag{222}$$

As before, the quantum fields are not Hermitian:

$$\hat{\Phi}_{\boldsymbol{k}}^\dagger = \hat{\Phi}_{-\boldsymbol{k}} \tag{223}$$

$$\hat{\Pi}_{\boldsymbol{k}}^\dagger = \hat{\Pi}_{-\boldsymbol{k}}. \tag{224}$$

2.7.2 The Hamiltonian of the Quantum Field

We canonically quantize our field Hamiltonian as

$$H = \int \mathrm{d}^D\boldsymbol{x} \frac{1}{2m}\pi_{\boldsymbol{x}}^2 + m\omega^2 \int \mathrm{d}^D\boldsymbol{x}\,(\nabla\varphi_{\boldsymbol{x}})^2 \tag{225}$$

$$\downarrow \text{canonically quantize}$$

$$\hat{H} = \int \mathrm{d}^D\boldsymbol{x} \frac{1}{2m}\hat{\pi}_{\boldsymbol{x}}^2 + m\omega^2 \int \mathrm{d}^D\boldsymbol{x}\,(\nabla\hat{\varphi}_{\boldsymbol{x}})^2 \tag{226}$$

and Fourier transforming using Eqs. 214 gives

$$\hat{H} = \int \mathrm{d}^D \boldsymbol{x} \int \mathrm{d}^D \boldsymbol{k} \int \mathrm{d}^D \boldsymbol{q} \left\{ \exp\left(-2\pi i \left(\boldsymbol{k}+\boldsymbol{q}\right)\cdot\boldsymbol{x}\right) \frac{1}{2m} \hat{\Pi}_{\boldsymbol{k}} \hat{\Pi}_{\boldsymbol{q}} + m\omega^2 \nabla \left(\exp\left(2\pi i \boldsymbol{k}\cdot\boldsymbol{x}\right) \hat{\Phi}_{\boldsymbol{k}} \right) \cdot \nabla \left(\exp\left(2\pi i \boldsymbol{q}\cdot\boldsymbol{x}\right) \hat{\Phi}_{\boldsymbol{q}} \right) \right\} \tag{227}$$

$$= \int \mathrm{d}^D \boldsymbol{x} \int \mathrm{d}^D \boldsymbol{k} \int \mathrm{d}^D \boldsymbol{q} \left\{ \exp\left(-2\pi i \left(\boldsymbol{k}+\boldsymbol{q}\right)\cdot\boldsymbol{x}\right) \frac{1}{2m} \hat{\Pi}_{\boldsymbol{k}} \hat{\Pi}_{\boldsymbol{q}} - \exp\left(2\pi i \left(\boldsymbol{k}+\boldsymbol{q}\right)\cdot\boldsymbol{x}\right) \left(2\pi\omega\right)^2 m \boldsymbol{k}\cdot\boldsymbol{q} \hat{\Phi}_{\boldsymbol{k}} \hat{\Phi}_{\boldsymbol{q}} \right\} \tag{228}$$

$$= \int \mathrm{d}^D \boldsymbol{k} \int \mathrm{d}^D \boldsymbol{q} \delta^D\left(\boldsymbol{k}+\boldsymbol{q}\right) \left\{ \frac{1}{2m} \hat{\Pi}_{\boldsymbol{k}} \hat{\Pi}_{\boldsymbol{q}} - \left(2\pi\omega\right)^2 m \boldsymbol{k}\cdot\boldsymbol{q} \hat{\Phi}_{\boldsymbol{k}} \hat{\Phi}_{\boldsymbol{q}} \right\} \tag{229}$$

$$= \int \mathrm{d}^D \boldsymbol{k} \left\{ \frac{1}{2m} \hat{\Pi}_{-\boldsymbol{k}} \hat{\Pi}_{\boldsymbol{k}} + \left(2\pi\omega\right)^2 m \boldsymbol{k}^2 \hat{\Phi}_{-\boldsymbol{k}} \hat{\Phi}_{\boldsymbol{k}} \right\} \tag{230}$$

$$= \int \mathrm{d}^D \boldsymbol{k} \left\{ \frac{1}{2m} \hat{\Pi}_{\boldsymbol{k}}^\dagger \hat{\Pi}_{\boldsymbol{k}} + \frac{1}{2} m\omega_{\boldsymbol{k}}^2 \hat{\Phi}_{\boldsymbol{k}}^\dagger \hat{\Phi}_{\boldsymbol{k}} \right\} \tag{231}$$

where we have defined

$$\omega_{\boldsymbol{k}} \triangleq \sqrt{8}\pi\omega \left|\boldsymbol{k}\right| \tag{232}$$

in the last line.

2.7.3 Creation and Annihilation Operators

The procedure now follows Section 2.5 exactly.
Defining the field creation operator

$$\hat{a}_{\boldsymbol{k}}^\dagger \triangleq \sqrt{\frac{m\omega_{\boldsymbol{k}}}{2}} \left(\hat{\Phi}_{-\boldsymbol{k}} - \frac{i}{m\omega_{\boldsymbol{k}}} \hat{\Pi}_{\boldsymbol{k}} \right) \tag{233}$$

we find

$$\left[\hat{a}_{\boldsymbol{k}}, \hat{a}_{\boldsymbol{q}}^\dagger\right] = \delta\left(\boldsymbol{k}-\boldsymbol{q}\right) \hat{\mathbb{I}} \tag{234}$$

and

$$\left[\hat{a}_{\boldsymbol{k}}, \hat{a}_{\boldsymbol{q}}\right] = \left[\hat{a}_{\boldsymbol{k}}^\dagger, \hat{a}_{\boldsymbol{q}}^\dagger\right] = 0. \tag{235}$$

We can now re-arrange to get

$$\hat{\Phi}_{\boldsymbol{k}} = \sqrt{\frac{1}{2m\omega_{\boldsymbol{k}}}} \left(\hat{a}_{\boldsymbol{k}} + \hat{a}_{-\boldsymbol{k}}^\dagger \right) \tag{236}$$

$$\hat{\Pi}_{\boldsymbol{k}} = i\sqrt{\frac{m\omega_{\boldsymbol{k}}}{2}} \left(\hat{a}_{\boldsymbol{k}}^\dagger - \hat{a}_{-\boldsymbol{k}} \right) \tag{237}$$

and so the Hamiltonian is

$$\hat{H} = \sum_{\boldsymbol{k}} \frac{1}{2m} \hat{\Pi}_{\boldsymbol{k}}^\dagger \hat{\Pi}_{\boldsymbol{k}} + \frac{1}{2} m\omega_{\boldsymbol{k}}^2 \hat{\Phi}_{\boldsymbol{k}}^\dagger \hat{\Phi}_{\boldsymbol{k}} \tag{238}$$

$$= \sum_{\boldsymbol{k}} \frac{\omega_{\boldsymbol{k}}}{4} \left(\hat{a}_{\boldsymbol{k}} - \hat{a}_{-\boldsymbol{k}}^\dagger \right) \left(\hat{a}_{\boldsymbol{k}}^\dagger - \hat{a}_{-\boldsymbol{k}} \right) + \frac{\omega_{\boldsymbol{k}}}{4} \left(\hat{a}_{\boldsymbol{k}}^\dagger + \hat{a}_{-\boldsymbol{k}} \right) \left(\hat{a}_{\boldsymbol{k}} + \hat{a}_{-\boldsymbol{k}}^\dagger \right) \tag{239}$$

$$= \sum_{\boldsymbol{k}} \frac{\omega_{\boldsymbol{k}}}{4} \left(\hat{a}_{\boldsymbol{k}}^\dagger \hat{a}_{\boldsymbol{k}} + \hat{a}_{\boldsymbol{k}} \hat{a}_{\boldsymbol{k}}^\dagger + \hat{a}_{-\boldsymbol{k}}^\dagger \hat{a}_{-\boldsymbol{k}} + \hat{a}_{-\boldsymbol{k}} \hat{a}_{-\boldsymbol{k}}^\dagger \right) \tag{240}$$

$$= \sum_{\boldsymbol{k}} \frac{\omega_{\boldsymbol{k}}}{4} \left(\hat{a}_{\boldsymbol{k}}^\dagger \hat{a}_{\boldsymbol{k}} + \left(\hat{a}_{\boldsymbol{k}}^\dagger \hat{a}_{\boldsymbol{k}} + \left[\hat{a}_{\boldsymbol{k}}, \hat{a}_{\boldsymbol{k}}^\dagger\right] \right) + \hat{a}_{-\boldsymbol{k}}^\dagger \hat{a}_{-\boldsymbol{k}} + \left(\hat{a}_{-\boldsymbol{k}} \hat{a}_{-\boldsymbol{k}}^\dagger + \left[\hat{a}_{-\boldsymbol{k}}, \hat{a}_{-\boldsymbol{k}}^\dagger\right] \right) \right) \tag{241}$$

$$= \sum_{\boldsymbol{k}} \frac{\omega_{\boldsymbol{k}}}{2} \left(\hat{a}_{\boldsymbol{k}}^\dagger \hat{a}_{\boldsymbol{k}} + \hat{a}_{-\boldsymbol{k}}^\dagger \hat{a}_{-\boldsymbol{k}} + \delta\left(0\right) \hat{\mathbb{I}} \right). \tag{242}$$

Unfortunately, by going to the continuum, the zero point energy has now become $\sum_{\boldsymbol{k}} \omega_{\boldsymbol{k}} \delta(0)/2$. This is a countably infinite sum over an uncountably infinite quantity! However, there is a standard method for dealing with this infinity, which we turn to now.

2.7.4 Normal Ordering

The infinity in Eq 242 is only the first of many infinities one will encounter in QFT. This one is perfectly harmless, since the absolute energy is meaningless: it is only differences in energies that have meaning. This infinity is also quite natural: we have placed a separate QHO at every position in space. Each QHO has a finite ground state energy. There are an uncountably infinite number of positions, and so the ground state energy must also be uncountably infinite.
We can remove this infinity by redefining our zero point energy (by an infinite amount).
There is a more systematic way to do this, however. Note that the infinity came about because we had to use the commutator $\left[\hat{a}_{\boldsymbol{k}}, \hat{a}^\dagger_{\boldsymbol{k}}\right]$. If we had written all creation operators to the left of all annihilation operators, we would not have needed to do this.
In fact, this infinity comes from an ambiguity introduced in canonical quantization: classical commuting observables become quantum non-commuting operators. Hence, the order can switch classically without causing a change, but quantum mechanically this changes the total energy. For example, if you look back at Eq 229, you will see that we did the $\boldsymbol{k}$ integral first, setting $\boldsymbol{k}$ to $-\boldsymbol{q}$, then we relabelled $\boldsymbol{k}$ to $\boldsymbol{q}$. You can instead do the $\boldsymbol{q}$ integral first, but this switches the order of the raising and lowering operators! Either choice should give the same physical results, and it is this ambiguity which is causing an apparent issue.
We define the *normal ordering* of a chain of operators to be such that all creation operators are to the left of all annihilation operators. We denote the normal ordering of an operator $\hat{\mathcal{O}}$ to be $:\hat{\mathcal{O}}:$. In our case

$$\hat{H} = \sum_{\boldsymbol{k}} \frac{\omega_{\boldsymbol{k}}}{4} \left(\hat{a}^\dagger_{\boldsymbol{k}} \hat{a}_{\boldsymbol{k}} + \hat{a}_{\boldsymbol{k}} \hat{a}^\dagger_{\boldsymbol{k}} + \hat{a}^\dagger_{-\boldsymbol{k}} \hat{a}_{-\boldsymbol{k}} + \hat{a}_{-\boldsymbol{k}} \hat{a}^\dagger_{-\boldsymbol{k}} \right) \tag{243}$$

and so its normal ordering, placing all creation operators to the left (without worrying about the resulting commutators) is

$$:\hat{H}: = \sum_{\boldsymbol{k}} \frac{\omega_{\boldsymbol{k}}}{2} \left(\hat{a}^\dagger_{\boldsymbol{k}} \hat{a}_{\boldsymbol{k}} + \hat{a}^\dagger_{-\boldsymbol{k}} \hat{a}_{-\boldsymbol{k}} \right). \tag{244}$$

This means that using normal ordered operators instead of the operators themselves, we subtract exactly the infinite term we had trouble with:

$$:\hat{H}: = \hat{H} - \sum_{\boldsymbol{k}} \frac{\omega_{\boldsymbol{k}}}{2} \delta(0) \hat{\mathbb{I}}. \tag{245}$$

In general, normal ordering will remove this particular type of infinity, introduced from the ordering ambiguity of quantum operators under canonical quantization. It is an infinite shift of the zero point energy, placing the ground state at zero energy.
Eq 244 is now a totally reasonable (normal-ordered) Hamiltonian, in which the $\boldsymbol{k}$ and $-\boldsymbol{k}$ modes are decoupled. As before, by noting the periodicity in $\boldsymbol{k}$ both terms now contribute an equal amount. Hence,

$$:\hat{H}: = \sum_{\boldsymbol{k}} \omega_{\boldsymbol{k}} \hat{a}^\dagger_{\boldsymbol{k}} \hat{a}_{\boldsymbol{k}}. \tag{246}$$

2.7.5 Excitations of a Quantum Field: Particles

Now let's take a step back and inspect our handywork, as we did for the quantum chain in Section 2.5.1 (you might like to re-read that section quickly).
Each operator $\hat{a}^\dagger_{\boldsymbol{k}}$ works exactly like a creation operator acts on a single QHO. Acting it on the ground state creates an excitation with a well defined wavevector. The ground state, or vacuum state, $|\Omega\rangle$ is a product of $N \to \infty$ independent modes. In 1D:

$$|\Omega\rangle = |0_{k=1}\rangle \otimes |0_{k=2}\rangle \otimes |0_{k=3}\rangle \otimes |0_{k=4}\rangle \otimes \dots \tag{247}$$

Note that, while positions have become continuous variables (forming an uncountably infinite set), the possible wavevectors remain countable. Our boundary conditions quantize the possible wavevectors. This is quite natural: a guitar string exists at an uncountably infinite number of positions between its two ends, yet its harmonics are standing waves which fit along the string. Hence the notation in Eq. 247 is acceptable.

Let's look at what happens when we act the field creation operator on the vacuum:

$$\hat{\Phi}^\dagger_{\boldsymbol{k}}|\Omega\rangle = \sqrt{\frac{1}{2m\omega_{\boldsymbol{k}}}}\left(\hat{a}^\dagger_{\boldsymbol{k}} + \hat{a}_{-\boldsymbol{k}}\right)|\Omega\rangle \tag{248}$$

$$= \frac{1}{\sqrt{2mE_{\boldsymbol{k}}}}\hat{a}^\dagger_{\boldsymbol{k}}|\Omega\rangle \tag{249}$$

$$= \frac{1}{2\sqrt{m}E_{\boldsymbol{k}}}\sqrt{2E_{\boldsymbol{k}}}|\boldsymbol{k}\rangle \tag{250}$$

$$= \frac{1}{2\sqrt{m}E_{\boldsymbol{k}}}|k\rangle. \tag{251}$$

In the last line we have adopted the standard notation

$$|k\rangle = \sqrt{2E_{\boldsymbol{k}}}|\boldsymbol{k}\rangle \tag{252}$$

which guarantees the Lorentz invariance of objects such as $\langle k|q\rangle = 2E_{\boldsymbol{k}}\delta\left(\boldsymbol{k}-\boldsymbol{q}\right)$ (a fact we established in Section 1.4). We are currently looking at a non-relativistic theory, but since QFT is so frequently applied in relativistic settings, notation such as Eq 252 has been established with this in mind.

What is $|k\rangle$ physically? It is nothing other than a *particle*! We call this particle a *phonon*. It has crystal-momentum $\boldsymbol{k}$, and energy $E_{\boldsymbol{k}} = \omega_{\boldsymbol{k}}$.

Note that Eq 249 means we can create a particle with momentum $\boldsymbol{k}$ using either $\hat{a}^\dagger_{\boldsymbol{k}}$ or $\hat{\Phi}^\dagger_{\boldsymbol{k}}$. The latter is more natural in a relativistic setting. In textbooks you are more likely to see particles created with the field operator than with the creation operator itself, although the reason is not always discussed.

At this point you might well ask why we would expect our theory to be relativistic, given that we started from balls and springs with absolute positions an and absolute times t. Essentially, the coarse-grained theory is relativistic even though the microscopic theory is not. We will return to this point in the next chapter.

We can denote multi-particle states as

$$|k_1, k_2\rangle = \hat{\Phi}^\dagger_{\boldsymbol{k}_2}\hat{\Phi}^\dagger_{\boldsymbol{k}_1}|\Omega\rangle \tag{253}$$

and so on. You can equally well create a particle at a specific position $\boldsymbol{x}$ by acting the real-space field creation operator on the vacuum:

$$\hat{\varphi}^\dagger_{\boldsymbol{x}}|\Omega\rangle = \int \mathrm{d}^D\boldsymbol{k}\exp\left(-2\pi i\boldsymbol{k}\cdot\boldsymbol{x}\right)\hat{\Phi}^\dagger_{\boldsymbol{k}}|\Omega\rangle \tag{254}$$

$$= \frac{1}{\sqrt{2m}}\int\frac{\mathrm{d}^D\boldsymbol{k}}{E_{\boldsymbol{k}}}\exp\left(-2\pi i\boldsymbol{k}\cdot\boldsymbol{x}\right)\sqrt{E_{\boldsymbol{k}}}|\boldsymbol{k}\rangle \tag{255}$$

$$= \frac{1}{\sqrt{2m}}\int \mathrm{d}^{D+1}k\,\delta\left(E^2 - m^2 - \boldsymbol{k}^2\right)\exp\left(-2\pi i\boldsymbol{k}\cdot\boldsymbol{x}\right)|k\rangle \tag{256}$$

which we conventionally denote

$$\hat{\varphi}^\dagger_{\boldsymbol{x}}|\Omega\rangle = |x\rangle. \tag{257}$$

Hence, acting the field creation operator at position $\boldsymbol{x}$ creates a particle at spacetime position x. Again, you can think by analogy of plucking a guitar string at a given position. The effect is to excite all the harmonics of the string, with different weights. In this case, 'plucking' means creating a particle from the vacuum, and the guitar string is the universe itself (here, a quantum crystal).

All of this intuition will carry across when you come to creating *photons* rather than *phonons*. However, photons are particles that do not require a medium through which to travel. This should again be natural, since photons are quantized particles of light, which requires no medium through which to propagate, while phonons are quantized particles of sound (lattice vibrations in a crystal) and sound does require a medium.

3 Path Integral Quantization

Until now you will mainly have seen quantum mechanics in terms of non-commuting operators and states. An alternative, equivalent approach, based on path integrals, was proposed by Dirac and formulated by Feynman. Dirac's observation was that the more familiar approach to QM puts space (an operator) and time (a variable) on different footing, so is hard to reconcile with relativity. Path integrals instead rely on the action S, which is a Lorentz scalar. Hence, it is the more natural approach to quantum mechanics in relativistic settings.
We will first quickly see path integral quantum mechanics, before generalising to quantum fields. This provides an alternative route to building quantum fields, equivalent to the canonical quantization method in Chapter 1.

3.1 Path Integral Quantum Mechanics

3.1.1 Motivation: Young's Slits

Consider a particle departing from position x_0 at time t_0, passing through one of two slits located at positions A and B at time t, and continuing onto a screen. The amplitude to find the particle at position x on the screen at time T, which we can denote $\psi(x,T)$, is the sum of the amplitudes to take each of the two paths. That is:

$$\begin{aligned}\psi(x,T) = &\,\mathrm{Amp}(x,T|A,t)\,\mathrm{Amp}(A,t|x_0,t_0)\\ &+\mathrm{Amp}(x,T|B,t)\,\mathrm{Amp}(B,t|x_0,t_0)\end{aligned} \tag{258}$$

where $\mathrm{Amp}(x,T|A,t)$ denotes the amplitude to find the particle at position x at time T given that it was certainly at position A at time t. Eq 258 shows that amplitudes in quantum mechanics play the role that probabilities play in classical systems. We then find quantum probabilities using

the **Born rule**:

$$\text{Probability} = |\text{Amplitude}|^2 . \tag{259}$$

What if there are not two slits to pass through at time t, but an infinite number of slits with nothing between them? In that case, the sum must be replaced by an integral:

$$\psi(x,T) = \int \mathrm{d}y \mathrm{Amp}(x,T|y,t)\,\mathrm{Amp}(y,t|x_0,t_0) . \tag{260}$$

But this should be true at all times between t_0 and T. Imagine 'slicing' time into N discrete segments. Then the amplitude to find a particle at position x at time T, given it started at x_0 at time t_0, must be:

$$\psi(x,T) = \int \mathrm{d}y_{N-1} \int \mathrm{d}y_{N-2} \ldots \int \mathrm{d}y_1 \mathrm{Amp}(x,T|y_{N-1},t_{N-1})\,\mathrm{Amp}(y_{N-1},t_{N-1}|y_{N-2},t_{N-2}) \ldots \mathrm{Amp}(y_1,t_1|x_0,t_0) . \tag{261}$$

Eq 261 encodes the idea that a quantum particle can be thought of as taking every possible path between two points. This is the basic idea behind path integral quantum mechanics. The object

$$\int \mathscr{D}\boldsymbol{y} \triangleq \prod_{n=1}^{N} \int \mathrm{d}y_n \tag{262}$$

which, in the limit $N \to \infty$ contains an infinite number of independent integrals, is called a 'functional integral' or 'path integral' (not to be confused with the 1D 'line integrals' seen in first year).

The posulates of path integral QM (Feynman, 1948)

(i) The Born rule: probability $= |\psi|^2$**.**

(ii) ψ **is given by a weighted sum over all possible classical trajectories** $\boldsymbol{x}(t)$**.**

(iii) in this sum, each path is weighted by $\exp(iS[\boldsymbol{x}])$**.**

3.1.2 The Propagator approach to QM

Single-particle non-relativistic QM amounts to solving the TDSE:

$$i\frac{\mathrm{d}|\psi\rangle}{\mathrm{d}t} = \hat{H}|\psi\rangle. \tag{263}$$

Assuming the Hamiltonian is time-independent we have the general solution

$$|\psi(t')\rangle = \exp\left(-i\hat{H}(t'-t)\right)|\psi(t)\rangle \tag{264}$$

which you can confirm by acting $i\hbar\mathrm{d}/\mathrm{d}t'$ on both sides. That is, if we have the solution at one time (which we obtain by solvign the TISE), we automatically have the solution at all subsequent times. It is convenient to define

the *time evolution operator*:

$$\hat{U}(t'-t) \triangleq \exp\left(-i\hat{H}(t'-t)\right) \tag{265}$$

where $\hat{U}$ is a unitary operator. We can project into the positon basis

$$\langle x'|\psi(t')\rangle = \langle x'|\hat{U}(t'-t)|\psi(t)\rangle \tag{266}$$

and use a

resolution of the identity (position basis):

$$\hat{\mathbb{I}} = \int \mathrm{d}x|x\rangle\langle x| \tag{267}$$

to find

$$\langle x'|\psi(t')\rangle = \int \mathrm{d}x\langle x'|\hat{U}(t'-t)|x\rangle\langle x|\psi(t)\rangle \tag{268}$$

or, in wavefunction notation,

$$\psi(x',t') = \int \mathrm{d}x K(x',t';x,t)\,\psi(x,t) \tag{269}$$

where we have defined

the propagator:

$$K(x',t';x,t) \triangleq \langle x'|\hat{U}(t'-t)|x\rangle. \tag{270}$$

That is, given a wavefunction at one point in space and time, the propagator 'propagates' the solution to any other point in space and time according to Eq 269.

3.1.3 The Propagator: free space

In free space we can calculate the propagator exactly.

$$K\left(x,t;x_0,t_0\right)=\langle x|\exp\left(-i\hat{H}\left(t-t_0\right)\right)|x_0\rangle. \tag{271}$$

In free space the potential $V=0$ and so

$$\hat{H}=\frac{\hat{p}^2}{2m}. \tag{272}$$

Therefore

$$K\left(x,t;x_0,t_0\right)=\langle x|\exp\left(-i\hat{p}^2\left(t-t_0\right)/\left(2m\right)\right)|x_0\rangle. \tag{273}$$

Inserting a

resolution of the identity (momentum basis):

$$\hat{\mathbb{I}}=\int_{-\infty}^{\infty}\mathrm{d}p|p\rangle\langle p| \tag{274}$$

gives

$$K\left(x,t;x_0,t_0\right)=\int_{-\infty}^{\infty}\mathrm{d}p\langle x|\exp\left(-i\hat{p}^2\left(t-t_0\right)/\left(2m\right)\right)|p\rangle\langle p|x_0\rangle. \tag{275}$$

By definition, $|p\rangle$ is an eigenstate of $\hat{p}$ with eigenvalue p:

$$\hat{p}|p\rangle=p|p\rangle \tag{276}$$

and so

$$K\left(x,t;x_0,t_0\right)=\int_{-\infty}^{\infty}\mathrm{d}p\langle x|\exp\left(-ip^2\left(t-t_0\right)/\left(2m\right)\right)|p\rangle\langle p|x_0\rangle \tag{277}$$

$$=\int_{-\infty}^{\infty}\mathrm{d}p\exp\left(-ip^2\left(t-t_0\right)/\left(2m\right)\right)\langle x|p\rangle\langle p|x_0\rangle \tag{278}$$

(where the momentum operator has become simply a variable). Recalling that momentum eigenstates are plane waves:

$$\langle x|p\rangle=\frac{1}{\sqrt{2\pi}}\exp\left(ipx\right) \tag{279}$$

We find

$$K\left(x,t;x_0,t_0\right)=\frac{1}{2\pi}\int_{-\infty}^{\infty}\mathrm{d}p\exp\left(i\left(p\left(x-x_0\right)-p^2\left(t-t_0\right)/\left(2m\right)\right)\right). \tag{280}$$

This is a Gaussian integral. It can be solved to give

the **propagator in free space**:

$$K\left(x,t;x_0,t_0\right)=\sqrt{\frac{m}{2\pi i}}\frac{1}{\sqrt{t-t_0}}\exp\left(\frac{im\left(x-x_0\right)^2}{2\left(t-t_0\right)}\right). \tag{281}$$

In free space, a particle can be thought as either taking a straight line between two points, or taking arbitrarily jagged paths between the same points. To see this, note that (for a free propagator only):

$$K\left(x_3,t_3;x_1,t_1\right)\equiv\int\mathrm{d}x_2K\left(x_3,t_3;x_2,t_2\right)K\left(x_2,t_2;x_1,t_1\right). \tag{282}$$

Proof:

$$\int \mathrm{d}x_2 K\left(x_3,t_3;x_2,t_2\right) K\left(x_2,t_2;x_1,t_1\right) = \int \mathrm{d}x_2 \langle x_3|\hat{U}\left(t_3-t_2\right)|x_2\rangle\langle x_2|\hat{U}\left(t_2-t_1\right)|x_1\rangle \tag{283}$$

remove the identity:

$$\begin{aligned}
\int \mathrm{d}x_2 K\left(x_3,t_3;x_2,t_2\right) K\left(x_2,t_2;x_1,t_1\right) &= \langle x_3|\hat{U}\left(t_3-t_2\right)\hat{U}\left(t_2-t_1\right)|x_1\rangle && (284)\\
&= \langle x_3|\exp\left(-i\hat{T}\left(t_3-t_2\right)\right)\exp\left(-i\hat{T}\left(t_2-t_1\right)\right)|x_1\rangle && (285)\\
&= \langle x_3|\exp\left(-i\hat{T}\left(t_3-t_1\right)\right)|x_1\rangle && (286)\\
&= K\left(x_3,t_3;x_1,t_1\right). && (287)
\end{aligned}$$

In this simplest case, we can break the propagator between two points into propagators between any intermediate number of points without introducing any approximation. The particle can either be thought of as following a straight line, or a sum over jagged lines.
When a potential $\hat{V}$ is introduced, we will be forced to sum over jagged lines. The free space calculation is simple because the only operator appearing in Eq 285 and Eq 286 is the kinetic operator $\hat{T}$: if a potential is present, $\left[\hat{T},\hat{V}\right]\neq 0$ and the exponentials cannot combine.

3.1.4 The propagator in a general potential: Deriving Feynman (iii)

In general,

$$\hat{H} = \hat{T}+\hat{V} = \frac{\hat{p}^2}{2m} + V\left(\hat{x}\right) \tag{288}$$

and so

$$K\left(x,t;x_0,t_0\right) = \langle x|\exp\left(-i\left(\hat{T}+\hat{V}\right)\left(t-t_0\right)\right)|x_0\rangle. \tag{289}$$

The problem with evaluating this is that $\hat{T}$ and $\hat{V}$ do not commute, and

$$\exp\left(A+B\right) \neq \exp\left(A\right)\exp\left(B\right) \tag{290}$$

unless $[A,B]=0$. Hence we cannot just insert an identity as in the case of a free particle.
The solution is to use *time slicing*. Break the path into tiny slices $\delta t = t/N$ with N large. Then,

$$\exp\left(-i\hat{H}t\right) = \left(\exp\left(-i\hat{H}\delta t\right)\right)^N \tag{291}$$

and

$$\exp\left(-i\left(\hat{T}+\hat{V}\right)\delta t\right) \approx \exp\left(-i\hat{T}\delta t\right)\exp\left(-i\hat{V}\delta t\right) \tag{292}$$

which becomes exact in the limit $\delta t\to 0$ (equivalently $N\to\infty$). Using time slicing we have:

$$K\left(x_N,t;x_0,t=0\right) = \langle x_N|\exp\left(-i\hat{H}\delta t\right)\times\ldots|x_0\rangle \tag{293}$$

where '…' signifies there are N identical copies within the bracket. Now we can insert $N-1$ resolutions of the identity into the position basis:

$$K\left(x_N,t;x_0,t=0\right) = \int \mathrm{d}x_{N-1}\ldots\int \mathrm{d}x_1 \langle x_N|\exp\left(-i\hat{H}\delta t\right)|x_{N-1}\rangle\ldots\langle x_1|\exp\left(-i\hat{H}\delta t\right)|x_0\rangle. \tag{294}$$

We must evaluate N integrals of the form

$$K_n \triangleq K\left(x_{n+1},\delta t;x_n,0\right) = \langle x_{n+1}|\exp\left(-i\hat{H}\delta t\right)|x_n\rangle = \langle x_{n+1}|\exp\left(-i\left(\hat{T}+\hat{V}\right)\delta t\right)|x_n\rangle. \tag{295}$$

We use the fact that the exponent is small to approximate

$$K_n \approx \langle x_{n+1}| \exp\left(-i\hat{T}\delta t\right) \exp\left(-i\hat{V}\delta t\right) |x_n\rangle \tag{296}$$

and we can now treat this slice like a free particle. That is, we insert a resolution of the identity into the momentum basis:

$$\begin{aligned} K_n &\approx \int \mathrm{d}p \langle x_{n+1}| \exp\left(-i\hat{T}\delta t\right) |p\rangle\langle p| \exp\left(-i\hat{V}\delta t\right) |x_n\rangle && (297)\\ &= \frac{1}{2\pi}\int \mathrm{d}p \exp\left(-ip^2\delta t/\left(2m\right)\right) \langle x_{n+1}|p\rangle \exp\left(-iV\left(x_n\right)\delta t\right) \langle p|x_n\rangle && (298)\\ &= \exp\left(-iV\left(x_n\right)\delta t\right) K_n^{\text{free}} && (299)\\ &= \exp\left(-iV\left(x_n\right)\delta t\right) \sqrt{\frac{m}{2\pi i\delta t}} \exp\left(im\left(x_{n+1}-x_n\right)^2/\left(2\delta t\right)\right). && (300) \end{aligned}$$

Putting it all together gives

$$K\left(x_N,t;x_0,0\right) = \lim_{N\to\infty}\left(\frac{m}{2\pi i\delta t}\right)^{N/2} \int \mathrm{d}x_{N-1}\ldots\int \mathrm{d}x_1 \exp\left(i\delta t\sum_{n=1}^{N-1}\left(\frac{m}{2}\left(\frac{x_{n+1}-x_n}{\delta t}\right)^2 - V\left(x_n\right)\right)\right). \tag{301}$$

The exponent in this limit simply becomes an integral:

$$\begin{aligned} \lim_{N\to\infty}\delta t\sum_{n=1}^{N-1}\left(\frac{m}{2}\left(\frac{x_{n+1}-x_n}{\delta t}\right)^2 - V\left(x_n\right)\right) &= \int_0^t \mathrm{d}t'\left(\frac{m}{2}\dot{x}^2 - V\left(x\right)\right) && (302)\\ &= \int_0^t \mathrm{d}t' L && (303)\\ &= S\left[x\right] && (304) \end{aligned}$$

where $S\left[x\right]$ is the action of the classical trajectory x. Therefore we have derived

$$K\left(x,t;x_0,t_0\right) = \int \mathscr{D}x \exp\left(iS\left[x\right]\right) \tag{305}$$

where we have defined the

the *functional integral* (aka *path integral*)

$$\int \mathscr{D}x \triangleq \lim_{N\to\infty}\prod_{n=1}^{N}\int \mathrm{d}x_n.$$

Note that the functional integral is a product of an infinite number of normal integrals(!). Eq. 305 is Feynman's postulate (iii): all classical paths, no matter how crazy, appear in the quantum sum over trajectories ('histories') with equal magnitude. All that changes is the complex phase assigned to each path.

3.1.5 Wick Rotation to Imaginary Time

Now that we have the basics of path integral QM, it is possible to see an otherwise unexpected connection to classical statistical mechanics (which will immediately become important). Note a resemblance between the time dependent Schroedinger equation of a free particle:

$$i\frac{\partial\psi}{\partial t} = -\frac{1}{2m}\frac{\partial^2\psi}{\partial x^2} \tag{306}$$

and the diffusion equation describing the classical evolution heat density $\rho\left(x,\tau\right)$ in position x and time τ:

$$\frac{\partial \rho}{\partial \tau} = D \frac{\partial^2 \rho}{\partial x^2} \tag{307}$$

where D is the diffusion coefficient. Formally, we can get from one to the other using a

Wick rotation to *imaginary time*:

$$t \to -i\tau. \tag{308}$$

If you ever read A Brief History of Time, and wondered what Hawking was on about when he kept referring to imaginary time, now you know!

The connection goes further. In QM we now know that the amplitude for a particle to start at x_i and end at x_f is

$$\langle x_f | \exp\left(-i\hat{H}t\right) |x_i\rangle = \int \mathscr{D}x \exp\left(iS\left[x\right]\right) \tag{309}$$

where

$$S\left[x\right] = \int_{t_i}^{t_f} \mathrm{d}t \left\{\frac{1}{2} m\dot{x}^2 - V\left(x\right)\right\}. \tag{310}$$

In statistical mechanics, we have a similar relation: the probability for a system to evolve from initial state x_i to final state x_f, at inverse temperature $\beta = 1/T$, is

$$\langle x_f | \exp\left(-\beta\hat{H}\right) |x_i\rangle = \int \mathscr{D}x \exp\left(-S_E\left[x\right]\right) \tag{311}$$

where

$$S_E\left[x\right] = \int_0^{\beta} \mathrm{d}\tau \left\{\frac{m}{2}\left(\frac{\mathrm{d}x}{\mathrm{d}\tau}\right)^2 + V\left(x\right)\right\}. \tag{312}$$

Here, S_E is called the 'Euclidean action'. This name derives from the fact that applying a Wick rotation to the Lorentz metric returns a Euclidean metric:

$$\text{Lorentz metric:} \quad x^2 - t^2 \tag{313}$$

$$t = -i\tau \downarrow \tag{314}$$

$$\text{Euclidean metric:} \quad x^2 + \tau^2. \tag{315}$$

A major use of Wick rotation is in making functional integrals converge. Whether a given path integral is well defined depends on the action. In quantum calculations, trajectories are weighted by complex phases $\exp\left(iS\right)$: convergence of the functional integrals requires the cancellation of rapidly varying phases for paths away from the paths of extremal action. In classical statistical mechanics, trajectories away from the extremal paths still contribute, but only with exponentially small probabilities $\exp\left(-S_E\right)$. These integrals are much more likely to converge. Hence, it is frequently a useful trick to Wick rotate to obtain convergence, then to analytically continue the solution back to real time.

3.1.6 Commutation Relations in Path Integral QM

Note that path integral QM uses the time slicing relation

$$\lim_{\delta t \to 0} \exp\left(-i\left(\hat{T} + \hat{V}\right)\delta t\right) = \lim_{\delta t \to 0} \exp\left(-i\hat{T}\delta t\right) \exp\left(-i\hat{V}\delta t\right) \tag{316}$$

suggesting

$$\left[\hat{T}, \hat{V}\right] = 0 \tag{317}$$

and therefore

$$[\hat{p}, \hat{x}] = 0. \tag{318}$$

Clearly, this cannot be the case if we are to reproduce quantum mechanics. In Lagrangian mechanics we deal with positions and velocities, rather than positions and momenta. Hence our question is whether x and $\dot{x}$ commute. We have:

$$[x, \dot{x}] = x(t+\delta t)\frac{x(t+\delta t)-x(t)}{\delta t} - x(t)\frac{x(t+\delta t)-x(t)}{\delta t} \tag{319}$$

$$= \left(\frac{x(t+\delta t)-x(t)}{\delta t}\right)^2 \delta t \tag{320}$$

$$= \dot{x}^2 \delta t. \tag{321}$$

In the first line we used the fact that

Operator ordering becomes time ordering in path integral QM.

At first glance this looks like it ought to be zero in the limit $\delta t \to 0$. But that might not be the case if $\dot{x}^2$ is infinite. Since all paths $x(t)$ are included in the functional integral, many (in fact, almost all) will be *nowhere differentiable*: they will have divergent $\dot{x}$ at all instants in time. So we have an infinity multiplying a zero.

To see what Eq 321 evaluates to, it is easiest to use Wick rotation. Consider the classical statistical mechanics problem of Brownian motion. Here, a pollen grain (say) receives random kicks from water molecules. Its change in velocity at each instant is random. But this does not mean its position at each instant is random, as its position at one instant must be close to where it was the instant before. Specifically, we know that for Brownian motion we have

$$\delta x^2 \propto \delta t \tag{322}$$

(as the pollen grain undergoes a random walk). This is formalised mathematically as Itô's lemma. Carrying this intuition back, we have that

$$[x, \dot{x}] = \dot{x}^2 \delta t = \left(\frac{\delta x}{\delta t}\right)^2 \delta t = \frac{\delta x^2}{\delta t} \tag{323}$$

and so, invoking Eq 322, we have

$$[x, \dot{x}] = 1. \tag{324}$$

That is, x and $\dot{x}$ do not commute, after all! Wick rotating back, and transforming from velocity to conjugate momentum, you obtain the usual canonical commutation relation

$$[x, p] = i. \tag{325}$$

The key point is this:

Non-commutation of operators arises in path integral QM from the jaggedness of the paths being summed over in the functional integral.

This will be important when we go to quantum fields, because it will allow us to sum over classical, commuting fields, but to arrive at non-commuting field operators.

3.2 Relativistic single-particle quantum mechanics

3.2.1 The Klein Gordon Equation

The Schroedinger equation you have seen in previous years is a single-particle non-relativistic equation. To make a relativistic quantum mechanics, we might try to simply canonically quantize the relativistic dispersion relation

$$E^2 - \boldsymbol{p}^2 - m^2 = 0 \tag{326}$$

using

$$E \to \hat{H} \triangleq i\partial_t$$
$$\boldsymbol{p} \to \hat{p} \triangleq -i\nabla.$$

The result is the

Klein Gordon Equation:

$$\left(\partial_t^2 - \nabla^2 + m^2\right)\varphi(x) = 0. \tag{327}$$

which can be written compactly using

$$\partial^2 \triangleq \partial^\mu \partial_\mu = \partial_t^2 - \nabla^2 \tag{328}$$

to give:

$$\left(\partial^2 + m^2\right)\varphi = 0. \tag{329}$$

Interestingly, Schroedinger devised the Klein Gordon equation before he devised the Schroedinger equation. He rejected it as an equation for the electron for various legitimate reasons. It is, however, a perfectly good quantum relativistic single-particle equation for spin-0 particles.
However, there are necessarily fundamental issues with any single-particle relativistic quantum theories, as we will now see.

3.2.2 The Failure of single-particle relativistic quantum theories

A question many students ask around this point is: can't we just upgrade the time evolution operator to a relativistic version, and create a single-particle relativistic quantum mechanics? Unfortunately this turns out not to work. Let's see why.
Consider the probability amplitude[4] $\mathcal{A}$ for a particle to propagate from position $\boldsymbol{x}$ at time t to position $\boldsymbol{x}'$ at time t':

$$\mathcal{A}(\boldsymbol{x}', t'; \boldsymbol{x}, t) = \langle \boldsymbol{x}' | \exp\left(-i\hat{H}(t'-t)\right) | \boldsymbol{x} \rangle. \tag{330}$$

We will use the relativistic Klein Gordon equation for the Hamiltonian. Inserting a complete set of momentum states as before,

$$\hat{\mathbb{I}} = \int \mathrm{d}^3\boldsymbol{p} |p\rangle\langle \boldsymbol{p}| \tag{331}$$

but this time using

$$\hat{H}|\boldsymbol{p}\rangle = \sqrt{\boldsymbol{p}^2 + m^2}|\boldsymbol{p}\rangle \tag{332}$$

gives

$$\begin{aligned}
\mathcal{A}(\boldsymbol{x}', t'; \boldsymbol{x}, t) &= \int \mathrm{d}^3\boldsymbol{p} \langle \boldsymbol{x}' | \exp\left(-i\hat{H}(t'-t)\right) |p\rangle\langle \boldsymbol{p}|\boldsymbol{x}\rangle && (333)\\
&= \int \mathrm{d}^3\boldsymbol{p} \langle \boldsymbol{x}' | \exp\left(-i\sqrt{\boldsymbol{p}^2 + m^2}(t'-t)\right) |p\rangle \exp(-i\boldsymbol{p}\cdot\boldsymbol{x}) && (334)\\
&= \int \mathrm{d}^3\boldsymbol{p} \langle \boldsymbol{x}'|p\rangle \exp\left(-i\sqrt{\boldsymbol{p}^2 + m^2}(t'-t)\right) \exp(-i\boldsymbol{p}\cdot\boldsymbol{x}) && (335)\\
&= \int \mathrm{d}^3\boldsymbol{p} \exp\left(-i\sqrt{\boldsymbol{p}^2 + m^2}(t'-t)/\hbar + i\boldsymbol{p}\cdot(\boldsymbol{x}'-\boldsymbol{x})\right). && (336)
\end{aligned}$$

A convenient co-ordinate choice is spherical polars with $\boldsymbol{p}\cdot(\boldsymbol{x}'-\boldsymbol{x}) = p|x'-x|\cos(\theta)$, where $p = |\boldsymbol{p}|$. This gives

[4] Note that this has dimensions of $\mathbb{L}^{-3}$, so the modulus square must be integrated over the start and end volumes to get a probability.

$$\mathcal{A}(\boldsymbol{x}', t'; \boldsymbol{x}, t) = 2\pi \int_0^\infty \mathrm{d}pp^2 \int_{-1}^1 \mathrm{d}(\cos\theta) \exp\left(-i\sqrt{p^2+m^2}\,(t'-t) + ip\,|\boldsymbol{x}'-\boldsymbol{x}|\cos\theta\right)$$
$$= \frac{4\pi}{|\boldsymbol{x}'-\boldsymbol{x}|} \int_0^\infty \mathrm{d}pp \exp\left(-i\sqrt{p^2+m^2}\,(t'-t)\right) \sin\left(p\,|\boldsymbol{x}'-\boldsymbol{x}|\right).$$

To proceed from here you can either do a contour integral[5], or do it the old-fashioned way: look it up in Gradshtein & Ryzhik. There you will find (7th edition equation 3.914.6):

$$\int_0^\infty p \exp\left(-\beta\sqrt{\gamma^2+p^2}\right) \sin(bp)\,\mathrm{d}p = \frac{b\beta\gamma^2}{\beta^2+b^2} K_2\left(\gamma\sqrt{\beta^2+b^2}\right) \tag{337}$$

where K_2 is a modified Bessel function of the second kind. This gives

$$\mathcal{A}(\boldsymbol{x}', t'; \boldsymbol{x}, t) = \frac{4\pi i\,(t'-t)\,m^2}{|\boldsymbol{x}'-\boldsymbol{x}|^2 - (t'-t)^2} K_2\left(m\sqrt{|\boldsymbol{x}'-\boldsymbol{x}|^2-(t'-t)^2}\right) \tag{338}$$

or

$$\mathcal{A}(\boldsymbol{x}', t'; \boldsymbol{x}, t) = \frac{4\pi i\,(t'-t)\,m^2}{\Delta s^2} K_2(m\Delta s) \tag{339}$$

where

$$\Delta s = \sqrt{|\boldsymbol{x}'-\boldsymbol{x}|^2 - (t'-t)^2} \tag{340}$$

is the Lorentz-invariant proper distance between the events. In the case where Δs is very large, *i.e.* well outside the light cone, we can look up the asymptotic expression for the Bessel function:

$$\lim_{x\to\infty} K_2(x) \sim \sqrt{\frac{\pi}{2x}} \exp(-x) \tag{341}$$

which gives

$$\lim_{\Delta s\to\infty} \mathcal{A}(\boldsymbol{x}', t'; \boldsymbol{x}, t) = i\frac{(t'-t)}{\Delta s}\left(\frac{2\pi m}{\Delta s}\right)^{3/2} \exp(-m\Delta s). \tag{342}$$

OK, so now say that the particle sets off from some volume $V = \mathrm{d}^3\boldsymbol{x}$ centred on $(\boldsymbol{x}, t)$, which is, for example, a box somewhere near Andromeda three seconds ago in your reference frame (therefore very much spacelike separated from you). The probability for you to find the particle in front of you now, within a volume $V' = \mathrm{d}^3\boldsymbol{x}'$, is

$$\text{Prob} = VV'\frac{(t'-t)^2}{\Delta s^2}\left(\frac{2\pi m}{\Delta s}\right)^3 \exp(-2m\Delta s). \tag{343}$$

The probability might be exponentially small... but it is non-zero!

This is the problem with single-particle relativistic quantum mechanics: there is always a finite probability to detect particles released from spacelike separations — that is, a finite probability to signal backwards in time. Since we've never seen anything signal backwards in time, we must reject any theory which says it is routine.

While this might seem like a problem which couldn't have come about before 1905, it was really known much earlier. For example, Michael Faraday took exception to the fact that waggling a magnet *here* can seemingly affect a magnet over *there*. Newton took exception to the fact that a planet *here* can seemingly affect an apple over *there*. The solution, in both cases, was to posit the existence of an invisible *field* which connects the observable objects and which is defined to act *locally*. Waggling the magnet *here* only affects the electromagnetic field *here*; a signal propagates through the field to the magnet over *there*.

Faced with action at a distance (particle propagation over spacelike separation), our solution is the same. We introduce quantum fields, which are defined to act locally.

[5] Most QFT textbooks seem to start the contour integral, but take some questionable steps and then only present the answer in a limit anyway.

3.3 Path Integral Field Quantization

The switch from single particles to fields is now carried out as follows:

$$\boldsymbol{x}(t) \to \varphi(\boldsymbol{x}, t) = \varphi(x^\mu). \tag{344}$$

That is, rather than $\boldsymbol{x}$ specifying the location of a single particle at time t, we instead have a field φ which can host any number of particles at any points in spacetime x^μ.

It is interesting to note that one can think of the position in the single-particle QM as a field in its own right (simply relabel $\boldsymbol{x}$ to $\boldsymbol{\varphi}$). In this sense, *QM is simply 0+1D QFT!*

It is both interesting and important to note that the fields in Eq 344 are classical, commuting fields! Nevertheless, the functional integral over these fields yields a quantum field theory. This work precisely as it did in single-particle path integral quantum mechanics in Section 3.1.6. The non-commutation of observables appears via the fractal nature of the fields being integrated over: a typical field in the integral is nowhere-differentiable.

We formulate a particular QFT by writing down its action, which is a functional of its (classical, commuting) fields. The action, as before, is the time integral of the Lagrangian:

$$S[\varphi] \triangleq \int \mathrm{d}t L[\varphi, \partial^\mu \varphi] \tag{345}$$

or the spacetime integral of the Lagrange density:

$$S[\varphi] = \int \mathrm{d}^4 x \mathscr{L}(\varphi, \partial^\mu \varphi). \tag{346}$$

Here, φ is a scalar field. This means it maps a number (well, 4-vector) x^μ to a scalar $\varphi(x^\mu)$. Standard notation you may see in textbooks denotes the space of x^μ as the 'base manifold' (recalling that Minkowski space is a manifold since it is locally $\mathbb{R}^N$), and the space of $\varphi(x^\mu)$ to be the 'target manifold'.

The particle excitations of scalar fields are spin-0 bosons. In the standard model the only example is the Higgs boson. But pions are composite spin-0 objects, and many particles in condensed matter physics (*e.g.* phonons) are spin-0.

We can also consider complex scalar fields with little extra work. The standard notation is to consider φ and φ^* to be independent fields. Then the action is a functional of both these fields:

$$\text{complex scalar field theory: } S[\varphi, \varphi^*]. \tag{347}$$

We will see that the particle excitations of φ^* can be interpreted as the antiparticles to those of φ.

The QFT governing scalar fields is called Klein Gordon (KG) theory. For real fields its Lagrange density is defined to be

$$\mathscr{L}_{\mathrm{KG}}(\varphi, \partial^\mu \varphi) \triangleq \frac{1}{2}\partial^\mu \varphi \partial_\mu \varphi - \frac{1}{2} m^2 \varphi^2. \tag{348}$$

Einstein summation notation is assumed, so there is an implicit sum over μ. For complex fields it is

$$\mathscr{L}^{\mathbb{C}}_{\mathrm{KG}}(\varphi, \varphi^*, \partial^\mu \varphi, \partial^\mu \varphi^*) \triangleq \partial^\mu \varphi^* \partial_\mu \varphi - m^2 \varphi^* \varphi. \tag{349}$$

3.4 Euler Lagrange Equations

In classical mechanics (using fields or otherwise) one constructs the action, considers all possible trajectories through phase space, and finds the trajectory that extremises the action subject to the initial and final boundary conditions.

In path integral quantum mechanics, one does the same, except non-extremal trajectories now contribute. The system evolves by taking all possible trajectories through phase space with each trajectory weighted by $\exp(iS)$. Still, the extremal trajectories are typically expected to give the biggest overall contribution, since the phase winding is slowest when S is near an extremum.

When we shift to QFT, trajectories (single-particle paths) become field configurations which can contain any number of particles. Otherwise the method is unchanged.

The sequences of field configurations which extremise the action again play a special role. In this case, they obey the single-particle quantum wave equation. Let's see how this works for the Klein Gordon field.

3.4.1 Example: the Klein Gordon field

The action defining the (real) Klein Gordon quantum field theory is

$$S_{\mathrm{KG}}\left[\varphi\right] \triangleq \int \mathrm{d}^4x \left(\frac{1}{2}\partial^\mu\varphi\partial_\mu\varphi - \frac{1}{2}m^2\varphi^2\right). \tag{350}$$

Here, $\varphi\left(x^\mu\right)$ is a real scalar field. Recall that, in the path integral formalism, fields are commuting: non-commutation appears from the fractal nature of typical field configurations in the functional integral. To find the associated Euler Lagrange equations we must extremise the action:

$$S_{\mathrm{KG}}\left[\varphi\right] = \int \mathrm{d}^4x \left(\frac{1}{2}\partial^\mu\varphi\partial_\mu\varphi - \frac{1}{2}m^2\varphi^2\right) \tag{351}$$

$$\downarrow \tag{352}$$

$$S_{\mathrm{KG}}\left[\varphi+\lambda\epsilon\right] = \int \mathrm{d}^4x \left(\frac{1}{2}\left(\partial^\mu\varphi+\lambda\partial^\mu\epsilon\right)\left(\partial_\mu\varphi+\lambda\partial_\mu\epsilon\right) - \frac{1}{2}m^2\left(\varphi+\lambda\epsilon\right)^2\right) \tag{353}$$

$$\downarrow \tag{354}$$

$$\frac{S_{\mathrm{KG}}\left[\varphi+\lambda\epsilon\right]}{\partial\lambda} = \int \mathrm{d}^4x \left(\frac{1}{2}\partial^\mu\epsilon\left(\partial_\mu\varphi+\lambda\partial_\mu\epsilon\right) + \frac{1}{2}\left(\partial^\mu\varphi+\lambda\partial^\mu\epsilon\right)\partial_\mu\epsilon - m^2\left(\varphi+\lambda\epsilon\right)\epsilon\right) \tag{355}$$

$$\downarrow \tag{356}$$

$$\left.\frac{S_{\mathrm{KG}}\left[\varphi+\lambda\epsilon\right]}{\partial\lambda}\right|_{\lambda=0} = \int \mathrm{d}^4x \left(\frac{1}{2}\partial^\mu\epsilon\partial_\mu\varphi + \frac{1}{2}\partial^\mu\varphi\partial_\mu\epsilon - m^2\varphi\epsilon\right). \tag{357}$$

We set this equal to zero. Integrating the first two terms by parts, using

$$\partial^\mu\partial_\mu\varphi = \partial_\mu\partial^\mu\varphi = \partial^2\varphi \tag{358}$$

gives

$$0 = \int \mathrm{d}^4x \left(-\partial^2\varphi - m^2\varphi\right)\epsilon \tag{359}$$

where we used the fact that $\epsilon\left(x^\mu\right)$ is defined to vanish at the limits of the spacetime integral. The only way Eq 359 can be true for any field ϵ is if the term in parentheses is zero. This term gives the Euler Lagrange equation corresponding to the Klein Gordon action:

$$\left(\partial^2 + m^2\right)\varphi = 0 \tag{360}$$

which is the Klein Gordon equation governing a single spin-0 relativistic particle.

3.5 Propagators & Green's Functions

3.5.1 Green's Functions for linear differential equations

Quantum mechanics is linear. In the 17th century George Green devised a general method for solving linear differential equations, which can all be written in the form

$$\hat{L}_x\psi\left(x\right) = f\left(x\right) \tag{361}$$

where $\hat{L}_x$ is some differential operator acting on x, and $f\left(x\right)$ is a source (forcing function) which renders the equation inhomogeneous. Working maximally generally like this, it might seem implausible that we could hope for a general solution. But looked at another way, in the finite dimensional case Eq 361 is just a matrix equation, and we know how to invert matrices to find solutions. Green's method generalises this intuition, providing a general solution to a linear differential equation. The principle is simple: if you can find the response a delta function source (an 'impulse'), you can integrate to get the solution for arbitrary sources.
We therefore begin by seeking to solve the special case

$$\hat{L}_xG\left(x,y\right) = \delta^D\left(x-y\right) \tag{362}$$

after which the general solution is given by

$$\psi(x) = \int \mathrm{d}^D y G(x,y) f(y). \tag{363}$$

To see this, just act $\hat{L}_x$ from the left:

$$\begin{aligned}\hat{L}_x \psi(x) &= \int \mathrm{d}^D y \hat{L}_x G(x,y) f(y) \\ &= \int \mathrm{d}^D y \delta(x-y) f(y) \\ &= f(x)\end{aligned}$$

which is the original equation including the general forcing term. The trick is that we can create any source $f(x)$ by linearly superposing delta functions:

$$f(x) = \int \mathrm{d}^D y f(y) \delta(x-y). \tag{364}$$

In cases of physical interest, symmetries generally constrain and simplify the problem. In particular, we often care about problems with translational symmetry, in which case

$$\text{translational symmetry} \implies G(x,y) = G(x-y). \tag{365}$$

3.5.2 Green's function for the Klein Gordon equation: momentum space

In practice this works as follows. Adding a delta function source to the Klein Gordon equation gives

$$\left(\partial^2 + m^2\right) G(x,y) = -i\delta^4(x-y) \tag{366}$$

where we have added a $-i$ in the definition purely for convenience (it neatens things later on). The Klein Gordon equation has spacetime translational symmetry, so this simplifies to

$$\left(\partial^2 + m^2\right) G(x-y) = -i\delta^4(x-y). \tag{367}$$

Further, translational symmetry suggests that we will do well to employ the Fourier transform:

$$G(x-y) = \int \frac{\mathrm{d}^4 p}{(2\pi)^4} \exp\left(-ip_\mu (x^\mu - y^\mu)\right) \tilde{G}(p) \tag{368}$$

to give

$$\begin{aligned}&\left(\partial^2 + m^2\right) \int \frac{\mathrm{d}^4 p}{(2\pi)^4} \exp\left(-ip_\mu (x^\mu - y^\mu)\right) \tilde{G}(p) = -i\delta^4(x) \\ &\int \frac{\mathrm{d}^4 p}{(2\pi)^4} \left(-p^2 + m^2\right) \exp\left(-ip_\mu (x^\mu - y^\mu)\right) \tilde{G}(p) = -i\delta^4(x) \\ &\int \frac{\mathrm{d}^4 p}{(2\pi)^4} \left(-p^2 + m^2\right) \exp\left(-ip_\mu (x^\mu - y^\mu)\right) \tilde{G}(p) = -i \int \frac{\mathrm{d}^4 p}{(2\pi)^4} \exp\left(-ip_\mu (x^\mu - y^\mu)\right)\end{aligned}$$

where we defined

$$p^2 \triangleq p_\mu p^\mu \tag{369}$$

and in the last line we used

$$\boxed{\int \frac{\mathrm{d}^4 p}{(2\pi)^4} \exp\left(-ip_\mu (x^\mu - y^\mu)\right) = \delta^4(x-y).} \tag{370}$$

Equating the integrands gives

the momentum-space Klein Gordon Green's function:

$$\tilde{G}(p) = \frac{i}{p^2 - m^2} = \frac{i}{E^2 - \boldsymbol{p}^2 - m^2}. \tag{371}$$

3.5.3 Klein Gordon Equation Green's function: real space

Using the Fourier transform in Eq 368 gives the real-space solution

$$G(x-y) = i \int \frac{\mathrm{d}^4 p}{(2\pi)^4} \frac{\exp\left(-ip_\mu \left(x^\mu - y^\mu\right)\right)}{p^2 - m^2}. \tag{372}$$

It is simple to check the this indeed obeys the Klein Gordon equation:

$$\left(\partial^2 + m^2\right) G(x-y) = i\left(\partial^2 + m^2\right) \int \frac{\mathrm{d}^4 p}{(2\pi)^4} \frac{\exp\left(-ip_\mu \left(x^\mu - y^\mu\right)\right)}{p^2 - m^2} \tag{373}$$

$$= i \int \frac{\mathrm{d}^4 p}{(2\pi)^4} \left(-p^2 + m^2\right) \frac{\exp\left(-ip_\mu \left(x^\mu - y^\mu\right)\right)}{p^2 - m^2} \tag{374}$$

$$= -i \int \frac{\mathrm{d}^4 p}{(2\pi)^4} \exp\left(-ip_\mu \left(x^\mu - y^\mu\right)\right) \tag{375}$$

$$= -i\delta^4(x-y). \tag{376}$$

To find an explicit formula for $G(x)$ (we set $y = 0$ for notational convenience), the Klein Gordon Green's function, we must use contour integration. Expanding the terms in Eq 372 gives

$$G(x) = i(2\pi)^{-4} \int \mathrm{d}E \mathrm{d}^3\boldsymbol{p} \frac{\exp\left(-iEt + i\boldsymbol{p}\cdot\boldsymbol{x}\right)}{E^2 - \boldsymbol{p}^2 - m^2} \tag{377}$$

$$= i(2\pi)^{-4} \int \mathrm{d}E \mathrm{d}^3\boldsymbol{p} \frac{\exp\left(-iEt + i\boldsymbol{p}\cdot\boldsymbol{x}\right)}{\left(E + E_{\boldsymbol{p}}\right)\left(E - E_{\boldsymbol{p}}\right)} \tag{378}$$

where

$$E_{\boldsymbol{p}} \triangleq \sqrt{\boldsymbol{p}^2 + m^2}. \tag{379}$$

Inspecting the energy integral, we find that it has simple poles at $E = \pm E_{\boldsymbol{p}}$. Since the integral is along the entire real axis, both poles are intercepted and cause divergences. To proceed, let us temporarily promote E to a complex variable

$$E \to z = z_1 + iz_2 \tag{380}$$

with $z_{1,2} \in \mathbb{R}$. Then

$$G(x) = i(2\pi)^{-4} \int_C \mathrm{d}z \mathrm{d}^3\boldsymbol{p} \frac{\exp(z_2 t)\exp\left(-iz_1 t + i\boldsymbol{p}\cdot\boldsymbol{x}\right)}{\left(z + E_{\boldsymbol{p}}\right)\left(z - E_{\boldsymbol{p}}\right)}. \tag{381}$$

If we take the contour C to be the real axis, this is exactly identical to the previous integral. However, for positive t we can connect $z = \infty + i0$ to $z = -\infty + i0$ by adding an infinite semicircular arc in the lower half plane: since $z_2 t$ is negative and infinite along this arc, this gives zero contribution to the integral. In this equivalent integral we now have a closed contour to which we can apply Cauchy's theorem. Similarly, if $t < 0$, we can add a contour which closes in the upper half plane.

We must then either deform the contour to miss the poles, or equivalently shift the poles to miss the contour. I will use the latter convention. There are four ways to move the poles (either can move up or down). Any choice gives a legitimate Green's function, but the forms differ. We will see the relationship between possible choices shortly.

For now, we will make the so-called 'Feynman prescription'. It is shown in Fig 4. We shift the pole at $-E_{\boldsymbol{p}}$ up infinitesimally to $-E_{\boldsymbol{p}} + i0^+$ (where 0^+ means an infinitesimal positive number) and the pole

Figure 4: The contour integral in Eq. 381, along with the Feynman prescription for handling the poles.

at $E_{\boldsymbol{p}}$ down to $E_{\boldsymbol{p}} - i0^+$. With this choice, only the $-E_{\boldsymbol{p}} + i0^+$ pole is enclosed when $t < 0$, and only the $E_{\boldsymbol{p}} - i0^+$ pole is enclosed when $t > 0$.
Overall, we have

$$G(x) = i(2\pi)^{-4} \oint_C \mathrm{d}z \mathrm{d}^3\boldsymbol{p} \frac{\exp(-iz_1 t + i\boldsymbol{p}\cdot\boldsymbol{x})}{(z + E_{\boldsymbol{p}} - i0^+)(z - E_{\boldsymbol{p}} + i0^+)} \exp(z_2 t) \tag{382}$$

where

$$C = \begin{cases} \infty \text{ clockwise semicircle in LHP}, & t > 0 \\ \infty \text{ anticlockwise semicircle in UHP}, & t < 0 \end{cases} \tag{383}$$

(the contour directions are fixed by the original integral, which runs from $-\infty$ to ∞).
Note that this choice of pole-shifts means that $E_{\boldsymbol{p}} \to E_{\boldsymbol{p}} - i0^+$ throughout; hence, Eq 382 can be compactly written as the Fourier transform of

$$\tilde{G}(p) = \frac{i}{p^2 - m^2 + i0^+} \tag{384}$$

which is a convenient way to remember it. Applying Cauchy's theorem (remembering that the pole encircled clockwise gives a negative contribution) gives

$$G(t > 0, \boldsymbol{x}) = (2\pi)^{-3} \int \frac{\mathrm{d}^3\boldsymbol{p}}{2E_{\boldsymbol{p}}} \exp(-iE_{\boldsymbol{p}} t + i\boldsymbol{p}\cdot\boldsymbol{x}) \tag{385}$$

$$G(t < 0, \boldsymbol{x}) = (2\pi)^{-3} \int \frac{\mathrm{d}^3\boldsymbol{p}}{2E_{\boldsymbol{p}}} \exp(iE_{\boldsymbol{p}} t + i\boldsymbol{p}\cdot\boldsymbol{x}) \tag{386}$$

Returning to the original argument $x - y$, this can be written as

$$G(x - y) = (2\pi)^{-3} \int \frac{\mathrm{d}^3\boldsymbol{p}}{2E_{\boldsymbol{p}}} \Theta\left(x^0 - y^0\right) \exp(i\,\boldsymbol{p}\cdot(\boldsymbol{x} - \boldsymbol{y})) \exp\left(-iE_{\boldsymbol{p}}\left|x^0 - y^0\right|\right). \tag{387}$$

where we have defined

the *Heaviside step function*

$$\Theta(x) = \begin{cases} 1, & x > 0 \\ 0, & x < 0 \end{cases} \tag{388}$$

and where x^0 is the time component of x^μ.

3.5.4 The Feynman Propagator

In Section we saw a formulation of single-particle non-relativistic QM in which the particle propagates from $(\boldsymbol{x}, t)$ to $(\boldsymbol{x}', t')$ via the propagator $K(\boldsymbol{x}', t'; \boldsymbol{x}, t)$. It is straightforward to construct a relativistic single-particle propagator $\Delta(x - y)$ using the Klein Gordon equation. In that case we require

$$\psi(x) = \int \mathrm{d}^4 x \Delta(x - y)\psi(y). \tag{389}$$

Back in the general theory of Green's functions in Section 3.5.1 we saw that the general solution to a linear equation

$$\hat{L}_x \psi(x) = f(x) \tag{390}$$

is given by

$$\psi(x) = \int \mathrm{d}^D y G(x, y) f(y).$$

Hence, the definition of the propagator tells us that it must be a Green's function for the Klein Gordon equation, with a 'source' which is the field (wavefunction, in the case of single particles) itself!
Returning to quantum fields, in which particles are created and destroyed, the propagator $\Delta(x-y)$ gives the amplitude to annihilate a particle at y and to create a particle at x. We are no longer restricted to the particle travelling from one event to the other. Instead, we integrate over all possible field configurations compatible with those two events.
Returning to the operator language of Section 2, we might formalise the stated behaviour of the propagator as follows:

$$\langle\Omega|\hat{\varphi}_{\boldsymbol{x}}\hat{\varphi}_{\boldsymbol{y}}|\Omega\rangle \tag{391}$$

which annihilates a particle at y and creates one at x. However, we need to ensure that the particle is not created before it is destroyed: we require $x^0 > y^0$. A simple nomenclature for this is to define

time ordering $\mathcal{T}$: earlier events are written to the right of later events.

This is not an operator; rather, it is a label indicating a convention that is being followed. Time ordering can be compared to the normal ordering used elsewhere. Hence, the propagator is really

$$\Delta(x-y) = \langle\Omega|\mathcal{T}\hat{\varphi}_{\boldsymbol{x}}\hat{\varphi}_{\boldsymbol{y}}|\Omega\rangle. \tag{392}$$

It is possible to use the formalism of Section 2 to verify directly that the Feynman propagator in Eq 392 is equal to the Klein Gordon propagator in Eq 387.
The Klein Gordon propagator $\Delta(x-y)$, using the Feynman prescription for placing the poles (Section 3.5.3), is termed the Feynman propagator. Feynman's choice of pole placements precisely produces the time ordering of the field operators in Eq 392.

3.6 Retarded and Advanced Green's Functions

Other choices of pole placements lead to alternative Green's functions for the Klein Gordon theory, with different physical meanings.
FINISH

3.7 Microcausality

3.7.1 Amplitudes versus signalling

We motivated the need for QFT by observing that relativistic single-particle propagators predict non-zero amplitudes to signal over spacelike separations. We did this by constructing a propagator using the single-particle Klein Gordon equation in Section 3.2.2.
We subsequently formulated the multi-particle Klein Gordon QFT, then we found its single-particle propagator in Eq 387. Has this actually helped? To answer this, we can again do the integral exactly. From Eq 387 we have

$$G(x) = (2\pi)^{-3}\int\frac{\mathrm{d}^3\boldsymbol{p}}{2E_{\boldsymbol{p}}}\Theta\left(x^0\right)\exp\left(i\,\boldsymbol{p}\cdot\boldsymbol{x}\right)\exp\left(-iE_{\boldsymbol{p}}\left|x^0\right|\right) \tag{393}$$

$$= (2\pi)^{-2}\int\frac{\mathrm{d}\left|\boldsymbol{p}\right|\mathrm{d}\left(\cos\left(\theta\right)\right)\boldsymbol{p}^2}{2E_{\boldsymbol{p}}}\Theta\left(x^0\right)\exp\left(i\left|\boldsymbol{p}\right|\left|\boldsymbol{x}\right|\cos\left(\theta\right)\right)\exp\left(-iE_{\boldsymbol{p}}\left|x^0\right|\right) \tag{394}$$

$$= \frac{i}{\left(2\pi\right)^2\left|\boldsymbol{x}\right|}\int\frac{\mathrm{d}\left|\boldsymbol{p}\right|\left|\boldsymbol{p}\right|}{E_{\boldsymbol{p}}}\Theta\left(x^0\right)\sin\left(\left|\boldsymbol{p}\right|\left|\boldsymbol{x}\right|\right)\exp\left(-iE_{\boldsymbol{p}}\left|x^0\right|\right) \tag{395}$$

on which we can use integration by parts (noting that $\left|\boldsymbol{p}\right|\exp\left(-iE_{\boldsymbol{p}}\left|x^0\right|\right)/E_{\boldsymbol{p}}$ happens to be an exact differential!) to give

$$G(x) = \frac{i}{\left(2\pi\right)^2\left|x^0\right|}\int\mathrm{d}\left|\boldsymbol{p}\right|\Theta\left(x^0\right)\cos\left(\left|\boldsymbol{p}\right|\left|\boldsymbol{x}\right|\right)\exp\left(-iE_{\boldsymbol{p}}\left|x^0\right|\right). \tag{396}$$

We can again look this up in Gradsteyn and Ryzhik (7th Edition, Eq 3.914.1) where we find

$$\int_0^\infty \mathrm{d}p \exp\left(-\beta\sqrt{\gamma^2+p^2}\right)\cos(bp) = \frac{\beta\gamma}{\sqrt{\beta^2+b^2}}K_1\left(\gamma\sqrt{\beta^2+b^2}\right) \tag{397}$$

and so

$$G(x-y) = \frac{1}{(2\pi)^2}\Theta\left(x^0-y^0\right)\frac{m}{\sqrt{|\boldsymbol{x}-\boldsymbol{y}|^2-(x_0-y_0)^2}}K_1\left(m\sqrt{|\boldsymbol{x}-\boldsymbol{y}|^2-(x_0-y_0)^2}\right) \tag{398}$$

$$= \frac{1}{(2\pi)^2}\Theta\left(x^0-y^0\right)\frac{m}{\Delta s}K_1(m\Delta s) \tag{399}$$

where

$$\Delta s = \sqrt{|\boldsymbol{x}-\boldsymbol{y}|^2-(x_0-y_0)^2} \tag{400}$$

is the proper length separating the events. As before, let's take the case $y=0$, with x very spacelike separated. Here we can use the asymptotic expression for K_1:

$$\lim_{x\to\infty} K_1(x) \sim \sqrt{\frac{\pi}{2x}}\exp(-x) \tag{401}$$

to find

$$\lim_{\Delta s\to\infty} G(x) \sim \frac{1}{(2\pi)^2}\frac{\sqrt{\pi m}}{\sqrt{2}\Delta s^{3/2}}\exp(-m\Delta s) \tag{402}$$

which, as before, is exponentially small but non-zero outside the light cone! So what has gone wrong? The problem lies with our interpretation. The single-particle propagator

$$G(x-y) = \langle\Omega|\mathcal{T}\hat{\varphi}_{\boldsymbol{y}}\hat{\varphi}_{\boldsymbol{x}}|\Omega\rangle \tag{403}$$

gives the amplitude to find the particle at x given it was at y, or equivalently to annihilate a particle at x and create one at y. In single-particle QM this would mean the particle travelled from x to y. But in QFT this is not true. A particle does not need to travel from x to y to be found at x at one instant, then y at the next. Rather, all particles are excitations of an underlying quantum field which existis everywhere and which can create particles anywhere.

The propagator in QFT is simply telling you that there is some correlation between the field amplitudes at x and y. That is perfectly natural. When I read the headline of my newspaper in the morning, my observation is perfectly correlated with that of my friend who reads the same newspaper at the same instant on the other side of the country, even though those events are spacelike separated. Similarly, the electron field exists across all of spacetime, and so there is a non-zero amplitude for you to find an electron in a box in front of you right now, and for your friend somewhere in the vicinity of Andromeda to find an electron in a box in front of them 3 seconds later. That doesn't allow you to signal between each other. You might even choose to interpret their electron to be your electron if you so wish, although since all electrons are identical that would be rather arbitrary.

What we do require is that we cannot signal at spacelike separation. This is encoded in QFT as the postulate of

> *microcausality*: operators $\hat{\mathcal{O}}_x$ and $\hat{\mathcal{O}}_y$ corresponding to spacelike-separated (bosonic) observables $(x-y)^2<0$ must commute:
>
> $$\left[\hat{\mathcal{O}}_x, \hat{\mathcal{O}}_y\right] = 0. \tag{404}$$

Hence, $G = \langle\Omega|\mathcal{T}\varphi_{\boldsymbol{y}}\varphi_{\boldsymbol{x}}|\Omega\rangle$ is simply not the correct object to consider when asking about signalling in QFT (although it is in single-particle QM). In QFT we instead need to consider

$$\langle\Omega|\left[\hat{\varphi}_{\boldsymbol{y}}, \hat{\varphi}_{\boldsymbol{x}}\right]|\Omega\rangle. \tag{405}$$

You can expand the field operators in terms of creation and annihilation operators (in the Heisenberg picture) to show explicitly that this vanishes for spacelike separations.

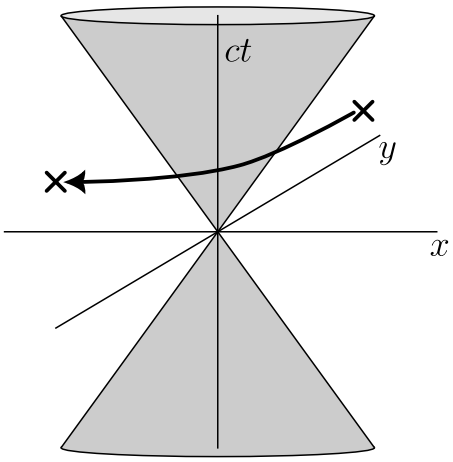


Figure 5: All points outside the lightcone are smoothly connected provided the spatial dimension is greater than 1.

However, there ought to be some more fundamental argument as to why QFTs disallow signalling. Peskin & Schroeder give a neat argument for real fields, as follows.
Define

$$D\left(x-y\right) \triangleq \langle\Omega|\hat{\varphi}_{\boldsymbol{x}}\hat{\varphi}_{\boldsymbol{y}}|\Omega\rangle \tag{406}$$

so that

$$\langle\Omega|\left[\hat{\varphi}_{\boldsymbol{x}},\hat{\varphi}_{\boldsymbol{y}}\right]|\Omega\rangle = D\left(x-y\right)-D\left(y-x\right). \tag{407}$$

In 3+1D, all spacelike points can be smoothly transformed into one another with an appropriate Lorentz transformation (see Fig 5). Hence, whenever $\left(x-y\right)^2<1$, we can always find some Lorentz transformation $\Lambda^{\mu}{}_{\nu}$ such that $\Lambda^{\mu}{}_{\nu}\left(x^{\nu}-y^{\nu}\right)=y^{\mu}-x^{\mu}$ (note that this is not possible for timelike separated events, as the timelike region contains two disconnected regions – past and future). This means that in some frame of reference $D\left(x-y\right)=D\left(y-x\right)$, and so $\langle\Omega|\left[\hat{\varphi}_{\boldsymbol{y}},\hat{\varphi}_{\boldsymbol{x}}\right]|\Omega\rangle=0$. But this is a Lorentz invariant quantity, and so must be independent of reference frame. Hence, $\langle\Omega|\left[\hat{\varphi}_{\boldsymbol{y}},\hat{\varphi}_{\boldsymbol{x}}\right]|\Omega\rangle=0$ whenever $\left(x-y\right)^2<1$. This is microcausality, and it is the statement of no signalling in QFT.

3.8 2-point functions

We saw in the previous section that the interpretation of the Green's function / propagator is a bit different in QFT compared to single-particle QM, since we no longer have to associate seeing a particle at x then at y with the particle propagating between the points. An alternative name for the object

$$G\left(x-y\right)=\langle\Omega|\mathcal{T}\hat{\varphi}^{\dagger}_{\boldsymbol{y}}\hat{\varphi}_{\boldsymbol{x}}|\Omega\rangle \tag{408}$$

is the '2-point function'. This emphasises that it really just gives the amplitude for particles to be found at two spacetime points. The nomenclature will prove helpful when we look at interacting quantum fields in the next section, where we will see general N-point functions (correlation amplitudes for multiple particles).
Returning to the path integral formalism, in which fields are commuting scalars but where quantum non-commutation arises from the jaggedness of the functional field integral, the 2-point function is given by

$$\langle\varphi_y\varphi_x\rangle \triangleq \frac{\int\mathscr{D}\varphi\varphi_y\varphi_x\exp\left(iS\left[\varphi\right]\right)}{\int\mathscr{D}\varphi\exp\left(iS\left[\varphi\right]\right)}. \tag{409}$$

There is a neat trick for calculating 2-point functions which we will turn to now.

3.8.1 The Partition Function

A key quantity of interest when it comes to practical QFT calculations is the *partition function*:

$$\mathscr{Z} \triangleq \int \mathscr{D}\varphi \exp\left(iS\left[\varphi\right]\right). \tag{410}$$

Recall that in statistical mechanics, while the partition function itself is a slightly abstract object, it is incredibly useful since all functions of state can be obtained from it (typically by taking derivatives). The same is true in QFT: all physical observables can be obtained from the partition function. This perhaps makes sense since the QFT is entirely specified by its action.
To proceed, let's look at a specific example.

3.8.2 The Klein Gordon Partition Function

The partition function for the real Klein Gordon field is

$$\mathscr{Z} = \int \mathscr{D}\varphi \exp\left(\frac{i}{2}\int \mathrm{d}^4x \left(\partial^\mu\varphi_x\partial_\mu\varphi_x - m^2\varphi_x^2\right)\right). \tag{411}$$

Integrating by parts we can rewrite this as

$$\mathscr{Z} = \int \mathscr{D}\varphi \exp\left(-\frac{i}{2}\int \mathrm{d}^4x\varphi_x\left(\partial^2 + m^2\right)\varphi_x\right). \tag{412}$$

We can Fourier transform to give

$$\mathscr{Z} = \int \mathscr{D}\varphi \exp\left(-\frac{i}{2}\int \frac{\mathrm{d}^{D+1}p}{(2\pi)^{D+1}}\tilde{\varphi}_p\left(-p^2 + m^2\right)\tilde{\varphi}_{-p}\right) \tag{413}$$

But recall that the Green's function of the Klein Gordon equation in Fourier space is

$$\tilde{G}_p = \frac{i}{p^2 - m^2} \tag{414}$$

and so

$$\mathscr{Z} = \int \mathscr{D}\varphi \exp\left(-\frac{1}{2}\int \frac{\mathrm{d}^{D+1}p}{(2\pi)^{D+1}}\tilde{\varphi}_p\tilde{G}_p^{-1}\tilde{\varphi}_{-p}\right). \tag{415}$$

We can return to real space, at least formally, with

$$\mathscr{Z} = \int \mathscr{D}\varphi \exp\left(-\frac{1}{2}\int \mathrm{d}^4x \int \mathrm{d}^4y\varphi_x G^{-1}_{x-y}\varphi_y\right) \tag{416}$$

where G_x^{-1} is the inverse[6] of the differential operator $i\left(\partial^2 + m^2\right)$ (we will neglect the hat on G for notational ease, and since it will always either be in either th position or momentum basis). This turns out to be a general rule for QFTs: the action for the free (non-interacting) field is simply two fields sandwiching the inverse of the Green's function.
Much of the utility of the partition function comes from the fact it is a Gaussian functional integral, and these are basically the only types of functional integrals we can do analytically. In fact it is this reason that we can do QFT at all! Let's remind ourselves about Gaussian integrals before putting this idea to use.

[6]The operator only has an inverse if it has no zero eigenvalues. This operator does have zero eigenvalues – but these are exactly what we learnt to deal with using the Feynman prescription.

3.8.3 Gaussian integrals

1D Gaussian integral
Recall the form of a 1D Gaussian integral:

$$I(a) = \int_{-\infty}^{\infty} \mathrm{d}x \exp\left(-\frac{1}{2}ax^2\right), \quad a \in \mathbb{C},\ \mathfrak{Re}(a) \geq 0. \tag{417}$$

This can be evaluated using an elegant trick, which is worth remembering:

$$I(a)^2 = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \mathrm{d}x\mathrm{d}y \exp\left(-\frac{1}{2}a\left(x^2+y^2\right)\right) \tag{418}$$

$$= \int_0^{2\pi} \mathrm{d}\theta \int_0^{\infty} \mathrm{d}r r \exp\left(-\frac{1}{2}ar^2\right) \tag{419}$$

$$= \frac{2\pi}{a} \tag{420}$$

and so

$$I(a) = \sqrt{\frac{2\pi}{a}}. \tag{421}$$

1D Gaussian integral with a source
Now consider

$$I(a,b) = \int_{-\infty}^{\infty} \mathrm{d}x \exp\left(-\frac{1}{2}ax^2 + bx\right). \tag{422}$$

This can be done by completing the square:

$$I(a,b) = \int_{-\infty}^{\infty} \mathrm{d}x \exp\left(-\frac{1}{2}a\left(x^2 - 2\frac{b}{a}x\right)\right) \tag{423}$$

$$= \int_{-\infty}^{\infty} \mathrm{d}x \exp\left(-\frac{1}{2}a\left(\left(x-\frac{b}{a}\right)^2 - \left(\frac{b}{a}\right)^2\right)\right) \tag{424}$$

$$= \exp\left(\frac{b^2}{2a}\right)\int_{-\infty}^{\infty} \mathrm{d}x \exp\left(-\frac{1}{2}a\left(x-\frac{b}{a}\right)^2\right) \tag{425}$$

$$= \exp\left(\frac{b^2}{2a}\right)\int_{-\infty}^{\infty} \mathrm{d}x \exp\left(-\frac{1}{2}ax^2\right) \tag{426}$$

$$= \sqrt{\frac{2\pi}{a}}\exp\left(\frac{b^2}{2a}\right) \tag{427}$$

where we changed integration variables to $x' = x - b/a$ and used the fact that the integral has an infinite range.
This calculation turns out to be very handy. If we denote the expectation value of x^n to be

$$\langle x^n \rangle = \int_{-\infty}^{\infty} \mathrm{d}x x^n \exp\left(-\frac{1}{2}ax^2 + bx\right) \tag{428}$$

then we have

$$\langle x^n \rangle = \left(\frac{\partial}{\partial b}\right)^n \int_{-\infty}^{\infty} \mathrm{d}x \exp\left(-\frac{1}{2}ax^2 + bx\right) \tag{429}$$

$$= \left(\frac{\partial}{\partial b}\right)^n \sqrt{\frac{2\pi}{a}}\exp\left(\frac{b^2}{2a}\right) \tag{430}$$

which is trivial to evaluate. We will shortly adopt this trick in QFT.

N-dimensional Gaussian integral
Now consider an N-dimenional Gaussian integral

$$I(A) = \int \mathrm{d}^N \boldsymbol{x} \exp\left(-\frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N} x_i^T A_{ij} x_j\right) \tag{431}$$

where A is a symmetric positive definite matrix (meaning all its eigenvalues are strictly positive). To do this integral, note that we can diagonalise A:

$$A = ODO^T \tag{432}$$

where O is an orthogonal matrix:

$$O^T O = \mathbb{I} \tag{433}$$

and

$$D = \mathrm{diag}\left(a_1, a_2, \ldots a_N\right). \tag{434}$$

Therefore, defining

$$\tilde{\boldsymbol{x}} \triangleq O^T \boldsymbol{x} \tag{435}$$

gives

$$\boldsymbol{x}^T A \boldsymbol{x} = \tilde{\boldsymbol{x}}^T D \tilde{\boldsymbol{x}} \tag{436}$$

$$= \sum_{i=1}^{N} a_i \tilde{x}_i^2. \tag{437}$$

In this diagonal basis the integral reduces to

$$I(A) = \prod_{i=1}^{N} \int \mathrm{d}\tilde{x}_i \exp\left(-\frac{1}{2} a_i \tilde{x}_i^2\right) \tag{438}$$

where the orthogonality condition, Eq 433, ensures that the Jacobian of the transformation (the equivalent of the r that appeared when switching to plane polar co-ordinates in 1D) is unity. Eq 438 is just a product of 1D Gaussian integrals, and so

$$I(A) = (2\pi)^{N/2} \left(a_1 a_2 \ldots a_N\right)^{-1/2}. \tag{439}$$

Recalling that the determinant of a matrix is the product of its eigenvalues, this gives

$$I(A) = \sqrt{\frac{(2\pi)^N}{\det(A)}}. \tag{440}$$

Gaussian Functional Integral
The limit $N \to \infty$ can be taken in the previous result without issue. Recalling that functions are infinite-dimensional vectors, and differential operators can be thought of as infinity-by-infinity matrices, we have

$$I\left(\hat{A}\right) = \int \mathscr{D}\varphi \exp\left(-\frac{1}{2}\int \mathrm{d}^{D+1}x \int \mathrm{d}^{D+1}y \varphi_x \hat{A}_{xy} \varphi_y\right) \tag{441}$$

$$= \sqrt{\frac{(2\pi)^\infty}{\det\left(\hat{A}\right)}} \tag{442}$$

where the determinant of an operator is again the product of its eigenvalues. The infinity does not prove to be a problem in practice, as we end up dividing throught by it later.

Notation: the symbol $\hat{A}_{xy}$ is intended to show the connection to the $N\times N$ matrix A_{ij} above. Operators that act locally (i.e. not at multiple spacetime events) must take the form

$$\hat{A}_{xy} = \delta^{D+1}(x-y)\,\hat{A}_x \tag{443}$$

so that

$$I\left(\hat{A}\right) = \int \mathscr{D}\varphi \exp\left(-\frac{1}{2}\int \mathrm{d}^{D+1}x \varphi_x \hat{A}_x \varphi_x\right). \tag{444}$$

It is sometimes convenient to use $\hat{A}_{xy}$, sometimes $\hat{A}_x$. We will use both below.

Gaussian Functional Integral with a source

As in 1D, we can introduce a source to our functional integral.

$$I[J] = \int \mathscr{D}\varphi \exp\left(-\frac{1}{2}\int \mathrm{d}^{D+1}x \int \mathrm{d}^{D+1}y \varphi_x \hat{A}_{xy}\varphi_y + \int \mathrm{d}^{D+1}x \varphi_x J_x\right). \tag{445}$$

Taking inspiration from 1D, we can make a change of variables

$$\varphi_x \to \varphi_x - \int \mathrm{d}^{D+1}z \hat{A}^{-1}_{xz} J_z \tag{446}$$

to give

$$I[J] = \int \mathscr{D}\varphi \exp\left(-\frac{1}{2}\int \mathrm{d}^{D+1}x \int \mathrm{d}^{D+1}y \left(\varphi_x - \int \mathrm{d}^{D+1}z \hat{A}^{-1}_{xz} J_z\right) \hat{A}_{xy} \left(\varphi_y - \int \mathrm{d}^{D+1}z \hat{A}^{-1}_{yz} J_z\right) + \frac{1}{2}\int \mathrm{d}^{D+1}x \int \mathrm{d}^{D+1}y J_{\tag{447}$$

$$= \exp\left(\frac{1}{2}\int \mathrm{d}^{D+1}x \int \mathrm{d}^{D+1}y J_x \hat{A}^{-1}_{xy} J_y\right) I[0]. \tag{448}$$

Hence

$$\int \mathscr{D}\varphi \exp\left(-\frac{1}{2}\int \mathrm{d}^{D+1}x \int \mathrm{d}^{D+1}y \varphi_x \hat{A}_{xy}\varphi_y + \int \mathrm{d}^{D+1}x \varphi_x J_x\right) = \exp\left(\frac{1}{2}\int \mathrm{d}^{D+1}x \int \mathrm{d}^{D+1}y J_x \hat{A}^{-1}_{xy} J_y\right) \sqrt{\frac{(2\pi)^\infty}{\det\left(\hat{A}\right)}} \tag{449}$$

which proves invaluable for calculating N-point functions.

3.9 2-point functions from the generating functional

The term $J(x)$ can be thought of as a fixed, specified function which acts as a current (source) of φ particles. Returning to our generic free field theory, the partition function with a source is now

$$\mathscr{Z}[J] = \int \mathscr{D}\varphi \exp\left(-\frac{1}{2}\int \mathrm{d}^{D+1}x \int \mathrm{d}^{D+1}y \varphi_x \hat{G}^{-1}_{xy}\varphi_y + i\int \mathrm{d}^{D+1}x \varphi_x J_x\right). \tag{450}$$

This is called the generating functional. Following the working above gives the exact result

$$\mathscr{Z}[J] = (2\pi)^{\infty/2}\sqrt{\det\left(\hat{G}\right)} \exp\left(-\frac{1}{2}\int \mathrm{d}^{D+1}x \int \mathrm{d}^{D+1}y J_x \hat{G}_{xy} J_y\right). \tag{451}$$

It is customary to drop the hat on the Green's function for notational convenience. The importance of this is that we can readily use it to generate N-point functions. For example, the two-point function (propagator) is given by

$$\langle \mathcal{T}\varphi_{x_1}\varphi_{x_2}\rangle = \mathscr{Z}^{-1}\left(-i\frac{\delta}{\delta J_{x_2}}\right)\left(-i\frac{\delta}{\delta J_{x_1}}\right)\mathscr{Z}[J]\Big|_{J=0}. \tag{452}$$

Let's check this explicitly. I will define

$$\varphi_1 \triangleq \varphi_{x_1} \tag{453}$$

$$G_{12} \triangleq G_{x_1 x_2} = G(x_1, x_2) \tag{454}$$

and so on, to save the notation getting too messy (this notation is used in some textbooks). Then

$$\mathscr{Z}^{-1}\left(-i\frac{\delta}{\delta J_2}\right)\left(-i\frac{\delta}{\delta J_1}\right)\mathscr{Z}[J] = \mathscr{Z}^{-1}\int \mathscr{D}\varphi \left(-i\frac{\delta}{\delta J_2}\right)\left(-i\frac{\delta}{\delta J_1}\right)\exp\left(-\frac{1}{2}\int \mathrm{d}^{D+1}x \int \mathrm{d}^{D+1}y \varphi_x G^{-1}_{xy}\varphi_y + i\int \mathrm{d}^{D+1}x \varphi_x \right. \tag{455}$$

$$= \mathscr{Z}^{-1}\int \mathscr{D}\varphi\varphi_1\varphi_2 \exp\left(-\frac{i}{2}\int \mathrm{d}^{D+1}x \int \mathrm{d}^{D+1}y \varphi_x G^{-1}_{xy}\varphi_y + i\int \mathrm{d}^{D+1}x \varphi_x J_x\right) \tag{456}$$

and so

$$\mathscr{Z}^{-1}\left(-i\frac{\delta}{\delta J_2}\right)\left(-i\frac{\delta}{\delta J_1}\right)\mathscr{Z}[J]\Big|_{J=0} = \frac{\int \mathscr{D}\varphi\varphi_1\varphi_2 \exp\left(-\frac{1}{2}\int \mathrm{d}^{D+1}x \int \mathrm{d}^{D+1}y \varphi_x G^{-1}_{xy}\varphi_y\right)}{\int \mathscr{D}\varphi \exp\left(-\frac{1}{2}\int \mathrm{d}^{D+1}x \int \mathrm{d}^{D+1}y \varphi_x G^{-1}_{xy}\varphi_y\right)} \tag{457}$$

$$= \frac{\int \mathscr{D}\varphi\varphi_2\varphi_1 \exp(iS_0)}{\int \mathscr{D}\varphi \exp(iS_0)} \tag{458}$$

as required. We can now calculate this using the exact solution in Eq 450:

$$\mathscr{Z}^{-1}\left(-i\frac{\delta}{\delta J_2}\right)\left(-i\frac{\delta}{\delta J_1}\right)\mathscr{Z}[J]\Big|_{J=0} = \mathscr{Z}^{-1}(2\pi)^{\infty/2}\sqrt{\det(G)}\left(-i\frac{\delta}{\delta J_2}\right)\left(-i\frac{\delta}{\delta J_1}\right)\exp\left(-\frac{1}{2}\int \mathrm{d}^{D+1}x \int \mathrm{d}^{D+1}y J_x G_{xy} J_y\right) \tag{459}$$

$$= \left(-i\frac{\delta}{\delta J_2}\right)\left(-i\frac{\delta}{\delta J_1}\right)\exp\left(-\frac{1}{2}\int \mathrm{d}^{D+1}x \int \mathrm{d}^{D+1}y J_x G_{xy} J_y\right)\Big|_{J=0} \tag{460}$$

$$= \left(-i\frac{\delta}{\delta J_2}\right)\left(\frac{i}{2}\int \mathrm{d}^{D+1}x J_x G_{x1} + \frac{i}{2}\int \mathrm{d}^{D+1}x J_x G_{1x}\right)\exp\left(-\frac{1}{2}\int \mathrm{d}^{D+1}x \int \mathrm{d}^{D+1}y J_{} \right. \tag{461}$$

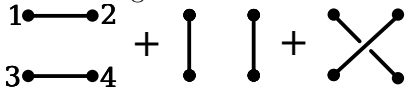
$$= \left(\frac{1}{2}(G_{12}+G_{21}) + \left(\frac{i}{2}\right)^2\left(\int \mathrm{d}^{D+1}x J_x (G_{x1}+G_{1x})\right)\left(\int \mathrm{d}^{D+1}x J_x (G_{x2}+G_{2x})\right)\right. \tag{462}$$

$$= G_{12} \tag{463}$$

noting in the last line that $G_{12} = G_{21}$. This confirms that the 2-point function is the Green's function (up to a normalisation).

3.10 Wick's Theorem

Another vital tool in QFT is Wick's theorem. To understand it, let's calculate some higher N-point functions from the generating functional. First, we can see that any $2N+1$-point function must vanish, since this involves an infinite integral over a product of an odd and even functional. This is simply the functional generalisation of the result



$$\int \mathrm{d}x x^{2n+1}\exp\left(-\frac{1}{2}Ax^2\right) = 0. \tag{464}$$

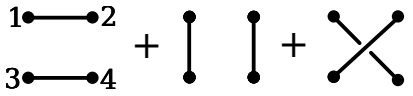


Figure 6: The 4-point function of a Gaussian field decomposes into all possible pairings of 2-point functions.

Let's look at the 4-point function, which corresponds (for example) to 2-particle scattering. Then

$$\langle \mathcal{T}\varphi_1\varphi_2\varphi_3\varphi_4\rangle = \mathscr{Z}^{-1}(-i)^4 \frac{\delta}{\delta J_1}\frac{\delta}{\delta J_2}\frac{\delta}{\delta J_3}\frac{\delta}{\delta J_4}\mathscr{Z}[J]\Big|_{J=0} \tag{465}$$

$$= (-i)^2 \frac{\delta}{\delta J_1}\frac{\delta}{\delta J_2}\left(G_{34} + i^2\left(\int \mathrm{d}^{D+1}x J_x G_{x3}\right)\left(\int \mathrm{d}^{D+1}x J_x G_{x4}\right)\right)\exp\left(-\frac{1}{2}\int \mathrm{d}^{D+1}x \int \mathrm{d}^{D+1}y J_x G_{xy} J_y\right)\Big| \tag{466}$$

$$= G_{12}G_{34} + G_{13}G_{24} + G_{14}G_{23}. \tag{467}$$

That is, the 4-point function is simply the sum of all possible products of pairs of 2-point functions! In general, we have

Wick's Theorem: the N-point function of a Gaussian theory is given by all possible products of pairings of 2-point functions.

Wick's theorem follows from the fact that the action is quadratic in the fields, and so our partition function is Gaussian. In statistics, you may have heard of this result as the fact that Normal (Gaussian) distributions are entirely characterised by their mean and variance. In the end, this is why QFT describes reality as harmonic oscillators at every point in spacetime: if it were anything other than a harmonic oscillator at each point, we would not be able to apply Wick's theorem.
A convenient representation of this pairing is shown in Fig 6. These are examples of Feynman diagrams, of which we will see much more in the next section.

4 Interacting quantum fields

The theories we have looked at so far are described by Gaussian partition functions or, equivalently, actions that are quadratic in the fields. These theories are trivial, in the sense that all N-point functions (which collectively describe everything there is to know) decompose identically into 2-point functions. To do useful calculations we need to consider interacting theories. The particles can either interact with themselves, or we can have multiple types of particle which can interact with one another.

4.1 φ^4 theory

We will focus on the first case: a scalar field with self-interactions, called φ^4 theory. The action is

$$S = S_0 + \int \mathrm{d}^{D+1}x \varphi_x J_x + S_{\text{int}} \tag{468}$$

where

$$S_0 \triangleq S_{\text{KG}} \tag{469}$$

$$S_{\text{int}} \triangleq \frac{\lambda}{4!}\int \mathrm{d}^{D+1}x \varphi_x^4. \tag{470}$$

You may hazard a guess as to where the name comes from! This action does not describe any particles in the standard model (nor does Klein Gordon), but it does describe other physical theories. For example, phonons emerging from balls and springs with an anharmonic potential between atoms, or the Ising model of interacting spins. The normalisation 1/4! turns out to be convenient later.

4.2 Perturbation Theory

To calculate the N-point functions for a non-Gaussian theory we can use perturbation theory. The basic idea is pretty simple. The generating functional is now given by

$$\mathscr{Z}_\lambda[J] = \int \mathscr{D}\varphi \exp\left(iS_0 + i\frac{\lambda}{4!}\int \mathrm{d}^{D+1}x\varphi_x^4 + i\int \mathrm{d}^{D+1}x\varphi_x J_x\right) \tag{471}$$

$$= \int \mathscr{D}\varphi \exp\left(iS_0 + i\int \mathrm{d}^{D+1}x\varphi_x J_x\right)\exp\left(i\frac{\lambda}{4!}\int \mathrm{d}^{D+1}x\varphi_x^4\right) \tag{472}$$

and we can expand the exponential of the interaction term as a Taylor series to give

$$\mathscr{Z}_\lambda[J] = \int \mathscr{D}\varphi \sum_{n=0}^{\infty}\frac{1}{n!}\left(i\frac{\lambda}{4!}\int \mathrm{d}^{D+1}x\varphi_x^4\right)^n \exp\left(iS_0 + i\int \mathrm{d}^{D+1}x\varphi_x J_x\right). \tag{473}$$

At each order in λ this is simply the expectation value of a polynomial of the fields, and any such object can be calculated using Wick's theorem. Let's take a look at some specific examples.

4.2.1 The interacting vacuum

The non-interacting field theory had an empty vacuum. But when interactions are present, the vacuum changes its form. To see this, we can calculate the 0-point function [7], which is simply $\mathscr{Z}_\lambda[J]$ itself. To $\mathcal{O}\left(\lambda^2\right)$ we have:

$$\mathscr{Z}_\lambda[J] = \int \mathscr{D}\varphi\left(1 + i\frac{\lambda}{4!}\int \mathrm{d}^{D+1}x\varphi_x^4 + \frac{1}{2}\left(i\frac{\lambda}{4!}\right)^2\int \mathrm{d}^{D+1}x\int \mathrm{d}^{D+1}y\varphi_x^4\varphi_y^4\right)\exp\left(iS_0 + i\int \mathrm{d}^{D+1}x\varphi_x J_x\right). \tag{474}$$

We work order-by-order in λ, and find

[7] I've not seen it called this, but it seems a sensible name to me!

$$\mathscr{Z}_\lambda[J] = \mathscr{Z}_0[J] + i\frac{\lambda}{4!}\int \mathrm{d}^{D+1}x \langle\varphi_x^4\rangle_0 + \frac{1}{2}\left(i\frac{\lambda}{4!}\right)^2 \int \mathrm{d}^{D+1}x \int \mathrm{d}^{D+1}y \langle\varphi_x^4\varphi_y^4\rangle_0 + \mathcal{O}\left(\lambda^3\right) \tag{475}$$

where we have introduced the convenient notation

$$\left\langle\hat{\mathcal{O}}\right\rangle_0 \triangleq \frac{\int \mathscr{D}\varphi \hat{\mathcal{O}} \exp\left(iS_0\right)}{\int \mathscr{D}\varphi \exp\left(iS_0\right)} \tag{476}$$

i.e. the familiar expectation value without the interaction term. We proceed using Wick's theorem, which says that any such $\left\langle\hat{\mathcal{O}}\right\rangle_0$ expectation value is equal to the sum of all possible products of 2-point functions. This gives

$$\mathscr{Z}_\lambda[J] = \mathscr{Z}_0[J] + i\frac{\lambda}{4!}\int \mathrm{d}^{D+1}x 3\langle\varphi_x^2\rangle_0\langle\varphi_x^2\rangle_0 \tag{477}$$

$$+ \frac{1}{2}\left(i\frac{\lambda}{4!}\right)^2 \int \mathrm{d}^{D+1}x \int \mathrm{d}^{D+1}y \left\{3\times 3\langle\varphi_x^2\rangle_0\langle\varphi_x^2\rangle_0\langle\varphi_y^2\rangle_0\langle\varphi_y^2\rangle_0 + 3\times 3\times 2\langle\varphi_x\varphi_y\rangle_0\langle\varphi_x\varphi_y\rangle_0\langle\varphi_x^2\rangle_0\langle\varphi_y^2\rangle_0\right\} \tag{478}$$

$$+ \mathcal{O}\left(\lambda^3\right). \tag{479}$$

The counting is as follows. In the $\mathcal{O}(\lambda)$ term, there are three ways for the first φ_x to pick a partner, and then the other pair is fixed. Similarly for the first $\mathcal{O}\left(\lambda^2\right)$ term (for x and y separately). For the second $\mathcal{O}\left(\lambda^2\right)$ term it is the same again for the xx and yy pairs, then there are two ways to choose the first xy pair and the second is fixed.
We now use the fact that the non-interacting 2-point function is the Green's function (multiplied by $1/2$) to give

$$\mathscr{Z}_\lambda[J] = \mathscr{Z}_0[J] + i\lambda\int \mathrm{d}^{D+1}x \frac{1}{8}G_{xx}^2 \tag{480}$$

$$+ \frac{1}{2}(i\lambda)^2 \int \mathrm{d}^{D+1}x \int \mathrm{d}^{D+1}y \left\{\frac{1}{8^2}G_{xx}^2G_{yy}^2 + \frac{1}{32}G_{xx}G_{xy}^2G_{yy} + G_{xy}^4\right\} \tag{481}$$

$$+ \mathcal{O}\left(\lambda^3\right). \tag{482}$$

4.2.2 Interacting 2-point function

Next we can redo the calculation of the 2-point function from Section 3.9. The $\mathcal{O}\left(\lambda^0\right)$ term is again the non-interacting expression. The $\mathcal{O}\left(\lambda^1\right)$ expression is

$$\langle\mathcal{T}\varphi_1\varphi_2\rangle = \mathcal{T}\mathscr{Z}_\lambda^{-1}\left(-i\frac{\delta}{\delta J_1}\right)\left(-i\frac{\delta}{\delta J_2}\right)\mathscr{Z}_\lambda[J]\Big|_{J=0} \tag{483}$$

$$= \mathcal{T}\mathscr{Z}_\lambda^{-1}\int\mathscr{D}\varphi\left(i\frac{\lambda}{4!}\int\mathrm{d}^{D+1}x\varphi_x^4\right)\left(-i\frac{\delta}{\delta J_1}\right)\left(-i\frac{\delta}{\delta J_2}\right)\exp\left(iS_0 + i\int\mathrm{d}^{D+1}x\varphi_xJ_x\right)\Big|_{J=0} \tag{484}$$

$$= \mathcal{T}\mathscr{Z}_\lambda^{-1}\int\mathscr{D}\varphi\left(i\frac{\lambda}{4!}\int\mathrm{d}^{D+1}x\varphi_x^4\right)(\varphi_1)(\varphi_2)\exp\left(iS_0 + i\int\mathrm{d}^{D+1}x\varphi_xJ_x\right)\Big|_{J=0} \tag{485}$$

$$= i\frac{\lambda}{4!}\int\mathrm{d}^{D+1}x\left\langle\mathcal{T}\varphi_x^4\varphi_1\varphi_2\right\rangle_0. \tag{486}$$

All possible contractions are shown schematically in Fig 7.
The result, counting all possibilities, is

$$\langle\mathcal{T}\varphi_1\varphi_2\rangle = i\frac{\lambda}{4!}\int\mathrm{d}^{D+1}x\left\{3\left\langle\varphi_x^2\right\rangle_0\left\langle\varphi_x^2\right\rangle_0\left\langle\varphi_1\varphi_2\right\rangle_0 + 4\times 3\left\langle\varphi_x\varphi_1\right\rangle_0\left\langle\varphi_x\varphi_2\right\rangle_0\left\langle\varphi_x^2\right\rangle_0\right\} \tag{487}$$

$$= i\lambda\int\mathrm{d}^{D+1}x\left\{\frac{1}{8}G_{xx}^2G_{12} + \frac{1}{2}G_{x1}G_{x2}G_{xx}\right\}. \tag{488}$$

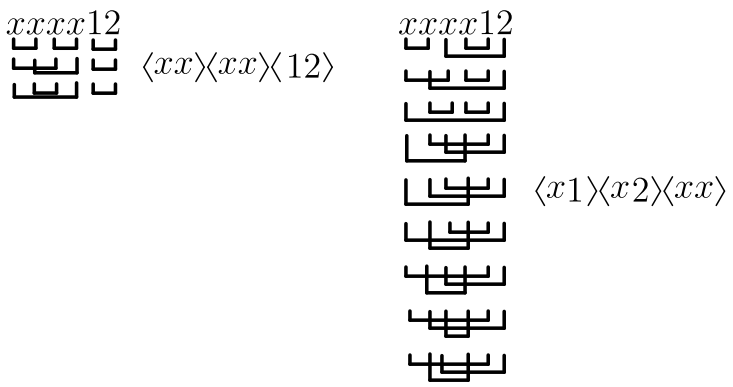


Figure 7: The possible pairs of terms when applying Wick's theorem to φ^4 theory to $\mathcal{O}(\lambda)$.

The counting of terms is as follows. For the $G_{xx}^2 G_{12}$ term, the first x has 3 ways to choose the second x with which it contracts. The remaining two xs are then forced. Hence, 3. For the $G_{x1}G_{x2}G_{xx}$ term, there are 4 choices of x for the 1 to pair with, leaving 3 choices for the 2 to pair with, then the remaining pair of xs must go together. Hence 4×3.

4.2.3 Interacting 4-point function

Since the 3-point function vanishes, the next simplest term is the 4-point function. At order λ^0 this is again the same as the unperturbed case. At order λ^1 it is given by

$$\langle \mathcal{T}\varphi_1\varphi_2\varphi_3\varphi_4\rangle_1 = i\frac{\lambda}{4!}\int \mathrm{d}^{D+1}x \left\langle \mathcal{T}\varphi_x^4\varphi_1\varphi_2\varphi_3\varphi_4\right\rangle_0 \tag{489}$$

$$= i\frac{\lambda}{4!}\int \mathrm{d}^{D+1}x \left\{4! G_{1x}G_{2x}G_{3x}G_{4x} + 3G_{xx}^2\left(G_{12}G_{34}+G_{13}G_{24}+G_{14}G_{23}\right)\right. \tag{490}$$

$$\left. + 4\times 3 G_{xx}\left(G_{1x}G_{2x}G_{34}+G_{1x}G_{3x}G_{24}+G_{1x}G_{4x}G_{23}+G_{2x}G_{3x}G_{14}+G_{2x}G_{4x}G_{13}+G_{3x}G_{4x}G_{12}\right)\right\} \tag{491}$$

$$= i\lambda\int \mathrm{d}^{D+1}x \left\{ G_{1x}G_{2x}G_{3x}G_{4x} + \frac{1}{8}G_{xx}^2\left(G_{12}G_{34}+2\text{ perms}\right) + \frac{1}{2}G_{xx}\left(G_{1x}G_{2x}G_{34}+5\text{ perms}\right)\right. \tag{492}$$

At higher orders of λ things quickly grow in combinatorial difficulty. There is a neat bookkeeping trick for dealing with this, to which we turn now.

4.3 Feynman Diagrams

4.3.1 Diagrams

Feynman invented his eponymous notation to simplify the rapidly growing complexity of the combinatorics just discussed.

The Feynman diagrams for the expansions in the previous sections appear in Fig 8 (the vacuum), Fig 9 (2-point function), and Fig 10 (4-point function).

Take the 2-point function. You start by drawing a vertex for each external field position (anything not integrated over, or equivalently anything explicitly labelled on the left): in our case, x_1 and x_2. Then draw a single point for each internal field position, in our case x. Now find all the ways to draw lines connecting these points, so that each external point meets precisely one line, and each internal point meets precisely 4 lines.

The reason each internal point has 4 lines meeting is that this is φ^4 theory. When you Taylor expand the action, every power of λ brings with it a product φ_x^4, so there is no other possibility for lines to meet. E.g. you can't have 2 lines meet at an internal point.

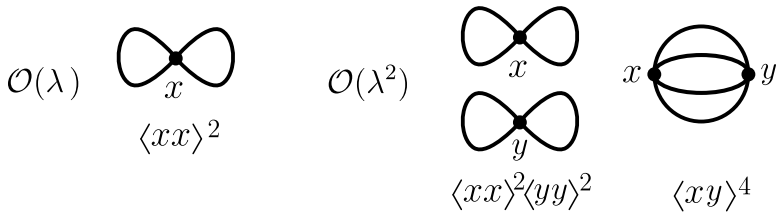


Figure 8: The Feynman diagrams for the vacuum of φ^4 theory to $\mathcal{O}(\lambda)$.

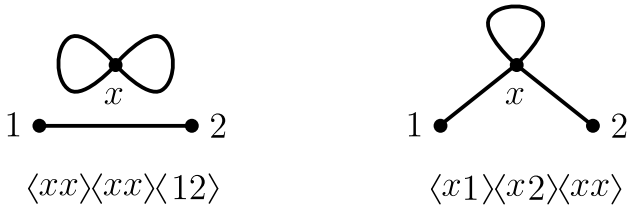


Figure 9: The Feynman diagrams for the 2-point functions in φ^4 theory to $\mathcal{O}(\lambda)$.

4.3.2 Symmetry factors

A remarkable result of the Feynman diagrams is that they instantly yield the difficult combinatorial prefactors. We term the 'symmetry factor' of a Feynman diagram the number the number of symmetry-equivalent diagrams it has. The prefactor of each term in the perturbative expansion is then given by the reciprocal of the diagram's symmetry factor. To see why this is, it's easiest to look at some examples.

In Eq 488 for the 2-point functions to order λ, the first term is $\frac{1}{8}G_{xx}^2 G_{12}$. Looking at the corresponding diagram in Fig 9 we see that the G_{12} line is fixed (x_1 and x_2 are fixed), but the G_{xx}^2 has some freedom. It has two lobes. We can flip the left lobe ($\times 2$), the right lobe ($\times 2$), and we can interchange the lobes ($\times 2$). The symmetry factor is $2 \times 2 \times 2 = 8$, and so the prefactor is $1/8$.

The second term is $\frac{1}{2}G_{x1}G_{x2}G_{xx}$. The corresponding diagram has one internal loop, which can be flipped ($\times 2$). Hence the diagram has a symmetry factor of 2, and the prefactor is $1/2$.

Now look at the 4-point functions in Eq 492, and the corresponding Feynman diagrams in Fig 10. The term $G_{1x}G_{2x}G_{3x}G_{4x}$ has a prefactor of 1. This is actually by construction: it is why we included the $1/4!$ factor in the definition of S_{int}. There are 4 ways for 1 to choose an x, 3 ways for the 2, 2 ways for the 3, and one way for the 4, giving 4! terms, cancelling the prefactor. Looking at the diagram, all legs connect to external events, so there is no freedom.

Other symmetry factors can be worked out using similar reasoning. Often it is quite tricky, in the general case; but practical QFTs such as QED are more restricted and typically have simpler symmetry factors.

4.3.3 Physical interpretation of Feynman diagrams

Feynman diagrams suggest a natural physical interpretation. Initially, it can be helpful to draw space and time axes (e.g. time heading up, space horizontal). Then you can interpret the lines (propagators) as worldlines of particles. Heading up, if lines meet (which must happen at a vertex), the particles annihilate. If lines emerge from a vertex, particles are created.

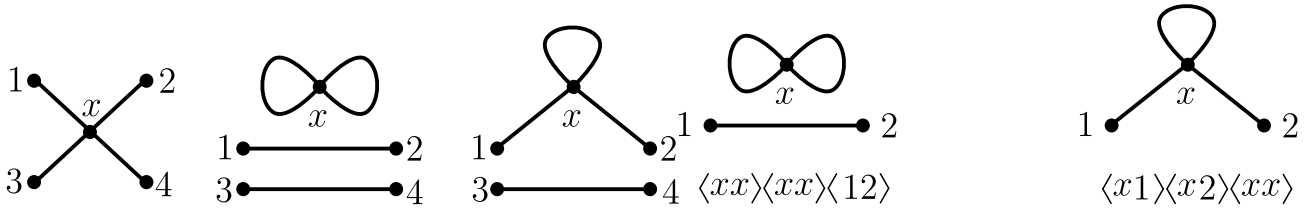


Figure 10: The Feynman diagrams for the 4-point functions in φ^4 theory to $\mathcal{O}(\lambda)$.

In φ^4 theory each vertex must be the meeting point of 4 lines, so if two head in, then two head out, and this is like two particles annihilating and another two being created at the same event.
The spacetime points of internal vertices are always integrated over: this means we create a quantum superposition of that event happening at any spacetime location.
Consider the 2-point function connecting events x_1 and x_2, which gives the probability amplitude to detect a particle at x_1 and at x_2 (recalling that this doesn't necessarily mean the same particle travels between those events – as the amplitude is non-zero for spacelike separation). In non-interacting φ^4 theory, there is an amplitude for the particle to propagate from x_1 to x_2. In the interacting theory, this is also possible. But inspecting Fig 9 (rightmost diagram), we see there is also an amplitude for the particle to start propagating from x_1, then to emit a particle at x, to catch that same particle *also at* x, then to propagate to x_2.
Giving a causal interpretation to intermediate events in Feynman diagrams is therefore tricky. Closed loops appear, and along any closed loop any spacetime axes must suggest a particle propagating forwards and backwards in time. That's how the intermediate particle can depart from x then arrive at x.
The spacetime axes are arbitrary: you could have drawn space heading up, time horizontal, and you would have found a different physical interpretation. This is a helpful reminder that the only measurable things occur at spacetime events corresponding to external vertices.

4.3.4 Feynman Rules

The Feynman diagrams are in one-to-one correspondence with the terms in the perturbation expansion of an interacting QFT. They are typically easy to draw, and it's fairly easy to check they've all been included at a given order. Hence, QFT typically amounts to drawing Feynman diagrams first, then converting them into their analytic expressions. This conversion is carried out by applying 'Feynman rules'.
For the case above, we have:

The Feynman rules for φ^4 theory (position space):

- To each propagator (line) assign G_{xy}
- To each vertex assign $i\lambda \int \mathrm{d}^{D+1}x$
- Divide by the symmetry factor.

In high-energy physics it is more common to work with the Fourier transformed propagators, since the boundary conditions typically take the form of specifying the momentum of a particle before and after a collision (recalling that the position and momentum cannot be simultaenously specified). In this case we have

The Feynman rules for φ^4 theory (momentum space):

- To each propagator (line) assign $\tilde{G}_p$
- To each vertex assign $i\lambda \left(2\pi\right)^4 \delta^4 \left(p_1 + p_2 + p_3 - p_4\right)$ (4-momentum conservation)
- integrate over internal momenta
- Divide by the symmetry factor.

4.3.5 The Connected Generating Functional

Inspecting and interpreting the Feynman diagrams, we see that many of them are rather trivial. For example, take the 2-point function G_{12}. At zeroth order this is the usual propagator. The first first-order correction $G_{12}G_{xx}^2$ is just this same propagator along with a 'bubble diagram', an event occuring elsewhere in spacetime (at x). Since fields were introduced to act locally, it would seem odd if we needed to include such unrelated spacetime events. And indeed we do not. What we are really interested in is *connected diagrams*: those in which the whole diagram is one connected piece.
To find the connected diagrams, we need to subtract off all the bubble diagrams. Note that every bubble appears in the vacuum: in fact, the vacuum is precisely the sum of all bubbles.

There is a very elegant trick for calculating only connected diagrams. Motivated by the observation that

$$\frac{\mathrm{d}}{\mathrm{d}x}\ln f = f^{-1}\frac{\mathrm{d}f}{\mathrm{d}x} \tag{493}$$

and the desire to divide out the vacuum, we replace our generating functional $\mathscr{Z}[J]$ with

The connected generating functional:

$$\mathscr{W}_\lambda[J] \triangleq \ln\left(\mathscr{Z}_\lambda[J]\right). \tag{494}$$

We use it precisely as we used $\mathscr{Z}_\lambda[J]$, but now we find only connected diagrams!
For example, consider the 2-point function to $\mathcal{O}(\lambda)$.
We have

$$\mathscr{W}_\lambda[J] = \ln\int\mathscr{D}\varphi\sum_{n=0}^{\infty}\frac{1}{n!}\left(i\frac{\lambda}{4!}\int\mathrm{d}^{D+1}x\varphi_x^4\right)^n\exp\left(iS_0 + i\int\mathrm{d}^{D+1}x\varphi_x J_x\right) \tag{495}$$

which to $\mathcal{O}(\lambda)$ is

$$\mathscr{W}_\lambda[J] = \ln\left[\mathscr{Z}_0[J] + i\frac{\lambda}{4!}\int\mathscr{D}\varphi\left(\int\mathrm{d}^{D+1}x\varphi_x^4\right)\exp\left(iS_0 + i\int\mathrm{d}^{D+1}x\varphi_x J_x\right)\right] \tag{496}$$

$$= \ln\left(\mathscr{Z}_0[J]\right) + \ln\left[1 + i\frac{\lambda}{4!}\mathscr{Z}_0^{-1}[J]\int\mathscr{D}\varphi\left(\int\mathrm{d}^{D+1}x\varphi_x^4\right)\exp\left(iS_0 + i\int\mathrm{d}^{D+1}x\varphi_x J_x\right)\right] \tag{497}$$

which can itself be Taylor expanded in λ, using

$$\ln(1+\lambda\epsilon) = \lambda\epsilon + \mathcal{O}\left(\lambda^2\right) \tag{498}$$

to give

$$\mathscr{W}_\lambda[J] = \ln\left(\mathscr{Z}_0[J]\right) + i\frac{\lambda}{4!}\mathscr{Z}_0^{-1}[J]\int\mathscr{D}\varphi\left(\int\mathrm{d}^{D+1}x\varphi_x^4\right)\exp\left(iS_0 + i\int\mathrm{d}^{D+1}x\varphi_x J_x\right). \tag{499}$$

After a reasonable bit of algebra (make sure to drop any terms you can spot will be zero along the way, e.g. $\langle\varphi_1\varphi_x^4\rangle_0$) you will find for the connected 2-point function

$$\langle\mathcal{T}\varphi_1\varphi_2\rangle_{\text{connected}} = \left(-i\frac{\delta}{\delta J_2}\right)\left(-i\frac{\delta}{\delta J_1}\right)\mathscr{W}_\lambda[J]\Big|_{J=0} \tag{500}$$

$$= \langle\mathcal{T}\varphi_1\varphi_2\rangle_0 + i\frac{\lambda}{4!}\int\mathrm{d}^{D+1}x\left(\langle\mathcal{T}\varphi_1\varphi_2\varphi_x^4\rangle_0 - \langle\mathcal{T}\varphi_1\varphi_2\rangle_0\langle\mathcal{T}\varphi_x^4\rangle_0\right). \tag{501}$$

The second term in parentheses subtracts off every bubble. The connected generating functional similarly leads to connected diagrams at all orders, and for all correlation functions.

4.4 Regularisation and Renormalisation

The point of all this perturbation theory is to find how interactions alter the 'bare', non-interacting N-point functions. Broadly, the answer is that the bare functions become renormalised. There are unfortunately two essentially totally unrelated meanings of the word 'renormalised', both introduced in QFT, and both of which are relevant here! Let's look at them in turn.

4.4.1 Renormalisation (Meaning I)

One of the main things we'd like to know about is the interacting 2-point function. The non-interacting 2-point function is the bare propagator for our theory. Adding interactions 'dresses' the propagator, changing its form. Often we will be interested to see what happens to one type of particle when we introduce a second type of particle with which it can interact. Typically we will expect that interactions will 'renormalise' the properties of both particles, meaning for example that their effective masses or charges change.

Figure 11:

A simple picture to hold in mind is to imagine trying to measure the mass of a tennis ball: you can do this by trying to accelerate the ball with a tennis racket, and measuring how much force it takes to impart a given acceleration. Now try the same experiment under water. It will take more force to impart that same acceleration, since the tennis ball is now interacting with the water molecules. Hence, your experiment tells you that the tennis ball's mass has been renormalised. This analogy also captures the fact that the renormalisation is accompanied by a change in the nature of the 'vacuum' itself, from air to water.
This idea applies wherever QFT is used. For example, the bare electron propagator in the standard model corresponds to a particle whose mass does not match the measured mass of the electron, because the electrons you can measure have already been renormalised by interactions with other fields (notably photons, but also W bosons, Higgs bosons, and anything that couples to electromagnetism or mass).
Often in condensed matter physics we are interested in precisely this type of scenario: two fields, with their associated particle types, where the particles of one field are much lighter than those of the other. In such cases, we typically carry out the functional integral of the lighter particles using perturbation theory, leading to a renormalised effective free (non-interacting) theory of the heavier particles. We call this process 'integrating out' a field, and it's a vital procedure in QFT. Let's look at a real example now.

4.4.2 The random phase approximation (Renormalisation I)

Consider the interaction of fast electrons with slow phonons in a crystal. Here, we have the QFT

$$S = S_0 + S_{\text{int}} \tag{502}$$

$$S_0 = \int \mathrm{d}^4 q \varphi_q D_q^{-1} \varphi_{-q} + \psi_k^\dagger G_k^{-1} \psi_k \tag{503}$$

$$S_{\text{int}} = g \int \mathrm{d}^4 q \varphi_q \psi_k^\dagger \psi_{k+q} \tag{504}$$

where $k = (\boldsymbol{k}, \omega)$ (with Euclidean metric),

$$D_q = \frac{-2\Omega_{\boldsymbol{q}}}{\omega^2 - \Omega_{\boldsymbol{q}}^2} \tag{505}$$

is the bare phonon propagator for phonons with energy $\Omega_{\boldsymbol{q}}$ at wavevector $\boldsymbol{q}$, and

$$G_k = \frac{1}{\omega - \xi_{\boldsymbol{k}}} \tag{506}$$

is the bare electron propagator, where $\xi_{\boldsymbol{k}}$ is the energy of an electron with crystal momentum $\boldsymbol{k}$. The energy of a particle is given by the locations of the poles in its Green's function. The positive energy pole of the bare phonon is therefore

$$\omega = \Omega_{\boldsymbol{q}}. \tag{507}$$

We can look at how the interactions with the electron field change this. If you try drawing connected Feynman diagrams to correct the phonon propagator, you will find that the lowest order diagram appears at order g^2. It is shown in Fig 11. However, there is a neat trick by which we can include an *infinite number* of Feynman diagrams of a similar form, also shown in that figure. We call this the random phase approximation or RPA (the name is for historical, and rather unconvincing, reasons).
It is important to note that the set of diagrams in the RPA is not all possible diagrams renormalising the phonon propagator. Rather, it is a mathematically convenient subset to consider, but there is no physical justification for discarding the other connected diagrams (the first of which appears at order g^4). RPA is an example of what is called an *uncontrolled approximation*: no effort is made to check it captures everything, and indeed it does not! This caveat in mind, let's proceed to calculate the RPA corrected phonon energy.

The loop which appears inside the propagator is called the polarisation bubble, or Lindhard function. It is

$$\chi_q = \int \mathrm{d}^4 k G_k G_{k+q}. \tag{508}$$

The frequency integral can be carried out using complex analysis, with the result

$$\chi_q = \sum_{\boldsymbol{k}} \frac{f\left(\xi_{\boldsymbol{k}+\boldsymbol{q}}\right) - f\left(\xi_{\boldsymbol{k}}\right)}{\omega + \xi_{\boldsymbol{k}+\boldsymbol{q}} - \xi_{\boldsymbol{k}}} \tag{509}$$

where

$$f\left(\xi\right) = \frac{1}{\exp\left(\frac{\xi}{T}\right) + 1} \tag{510}$$

is the Fermi Dirac distribution. The $\boldsymbol{k}$ integral has been replaced by a discrete sum (since this is a crystal). The sum requires $\xi_{\boldsymbol{k}}$ to be specified in order to be evaluated, and this is generally done numerically. The interacting phonon propagator is then

$$D_q^{\mathrm{RPA}} = D_q + D_q\left(g^2\chi_q\right)D_q + D_q\left(g^2\chi_q\right)D_q\left(g^2\chi_q\right)D_q + D_q\left(g^2\chi_q\right)D_q\left(g^2\chi_q\right)D_q\left(g^2\chi_q\right)D_q + \ldots \tag{511}$$

This general form is called a Dyson series. It is simply a geometric series of Feynman diagrams. As such, the infinite sum can be carried out exactly, by noticing that

$$\begin{aligned} D_q^{\mathrm{RPA}} &= D_q + D_q\left(g^2\chi_q\right)\left(D_q + D_q\left(g^2\chi_q\right)D_q + D_q\left(g^2\chi_q\right)D_q\left(g^2\chi_q\right)D_q + \ldots\right) && (512)\\ &= D_q + D_q\left(g^2\chi_q\right)D_q^{\mathrm{RPA}} && (513)\end{aligned}$$

and so, subtracting the second term and gathering things together:

$$\left(1 - D_q g^2 \chi_q\right) D_q^{\mathrm{RPA}} = D_q \tag{514}$$

giving the final result

$$\begin{aligned} D_q^{\mathrm{RPA}} &= \left(1 - D_q g^2 \chi_q\right)^{-1} D_q && (515)\\ &= \frac{-2\Omega_{\boldsymbol{q}}}{\omega^2 - \Omega_{\boldsymbol{q}}^2 + 2g^2\Omega_{\boldsymbol{q}}\chi_q}. && (516)\end{aligned}$$

The positive pole of this renormalised propagator occurs at

$$\omega = \sqrt{\Omega_{\boldsymbol{q}}^2 - 2g^2\Omega_{\boldsymbol{q}}\chi_q}. \tag{517}$$

That is, by interacting with the electrons, the energy ω of the phonon at a given wavevector has decreased.

4.4.3 1PI Diagrams and Self-Energies

The Dyson series used in the RPA diagram summation is so useful that it has become a standard method of organising Feynman diagrams in QFT. We define a '1-Particle Irreducible' (1PI) diagram to be:

1PI diagram: any Feynman diagram which remains fully connected after cutting any one internal line.

With this definition, we can always decompose the infinite set of Feynman diagrams into a Dyson series, as in Fig 12. How do we know this is always possible? It is by definition! The Dyson series is defined to be the individual 1PI parts connected together with single edges (cutting any one of which will disconnect the diagram). So the Dyson series captures everything that is not 1PI, and since we're summing everything, the bits in the 1PI blobs have to be all the bits of diagrams which remain

Figure 12: Any infinite set of Feynman diagrams can be decomposed into a Dyson series of 1PI diagrams.

connected after a single cut. You can think of 1PI diagrams as a bit like the prime numbers of Feynman diagrams: just as any integer can be factored into a product of primes, any Feynman diagram can be factored into a product of 1PI diagrams.
As in the RPA, we can now deduce the renormalisation of our propagator (or any other N-point function) by an infinite Dyson series of 1PI diagrams. When renormalising a propagator we define the 1PI blob to be a 'self energy' Σ. Then, as in RPA, we can argue that

$$G^{\text{ren}} = G + G\Sigma G^{\text{ren}} \tag{518}$$

$$\downarrow$$

$$G^{\text{ren}} = (1 - G\Sigma)^{-1} G. \tag{519}$$

Example: non-relativistic electron propagator
The clearest example of this effect is with the (spinless, non-relativistic) electron propagator which is sometimes used in condensed matter physics. Here we have

$$G = \frac{1}{\omega - \xi_{\boldsymbol{k}} + i\epsilon} \tag{520}$$

where ϵ is an infinitesimal positive real number (specifying how the pole shifts off the real axis in contour integrals). Ignoring ϵ, the Green's function has a pole at $\omega = \xi_{\boldsymbol{k}}$, showing that the energy ω of the electron with crystal momentum $\boldsymbol{k}$ is $\xi_{\boldsymbol{k}}$.
The probability to find the electron with wavevector $\boldsymbol{k}$ and energy ω is given by the 'spectral function'

$$A(\boldsymbol{k}, \omega) \triangleq -\frac{1}{\pi} \mathfrak{Im} G(\boldsymbol{k}, \omega) \tag{521}$$

$$= \frac{1}{\pi} \frac{\epsilon}{(\omega - \xi_{\boldsymbol{k}})^2 + \epsilon^2}. \tag{522}$$

Noting that this is simply a Lorentzian of width ϵ we see that

$$\lim_{\epsilon \to 0^+} A(\boldsymbol{k}, \omega) = \delta(\omega - \xi_{\boldsymbol{k}}) \tag{523}$$

which says that the electron has 100% probability to be found with energy $\omega = \xi_{\boldsymbol{k}}$. As a quick check that this makes sense, we can confirm that the electron must exist at *some* energy, and so

$$\int \mathrm{d}\omega A(\boldsymbol{k}, \omega) = 1 \tag{524}$$

which you can confirm by doing the Lorentzian integral. (This is called a 'sum rule').
What happens when the electron propagator is renormalised by an infinite Dyson series? We find

$$G^{\text{ren}} = (1 - G\Sigma)^{-1} G \tag{525}$$

$$= \frac{1}{\omega - \xi_{\boldsymbol{k}} - \Sigma + i\epsilon}. \tag{526}$$

Writing

$$\Sigma \triangleq \Sigma' + i\Sigma'' \tag{527}$$

for real Σ',Σ'', the spectral function becomes

$$A(\boldsymbol{k},\omega) = -\frac{1}{\pi}\mathfrak{Im}G(\boldsymbol{k},\omega) \tag{528}$$

$$= \frac{1}{\pi}\frac{\Sigma'' + \epsilon}{(\omega - \xi_{\boldsymbol{k}} - \Sigma')^2 + (\epsilon - \Sigma'')^2} \tag{529}$$

where we can now drop the $\epsilon \ll \Sigma''$ to give

$$A(\boldsymbol{k},\omega) = \frac{1}{\pi}\frac{\Sigma''}{(\omega - \xi_{\boldsymbol{k}} - \Sigma')^2 + \Sigma''^2}. \tag{530}$$

Two things have happened. First, the real part of the self energy Σ' has shifted the peak of A to $\omega = \xi_{\boldsymbol{k}} + \Sigma'$, renormalising the energy. Second, the imaginary part Σ'' has given a finite width to the spectral function Lorentzian. The electron of wavevector $\boldsymbol{k}$ is no longer certain to be at the peak in A; rather, it can now have any energy (with the probability for it to have energy between ω and $\omega + \mathrm{d}\omega$ given by $A\mathrm{d}\omega$).
Note that the sum rule is still obeyed: the total probability to find the electron with some energy is still 1.

4.4.4 Renormalisation II

In φ^4 theory, the field interacts with itself. This self-interaction will renormalise (I) the N-point functions. To order λ, with connected diagrams, the propagator is renormalised as

$$G_{12} \to G_{12} + i\frac{\lambda}{2}\int \mathrm{d}^4x G_{1x}G_{xx}G_{x2} \tag{531}$$

This time, however, there is a problem. The term

$$G_{xx} = G(x - x) = G(0) \propto \delta^4(0) \tag{532}$$

is infinite. We saw similar infinities before, which we eliminated by requiring operators to be normal ordered. This new type of infinity is sometimes called an 'ultraviolet divergence', since it has occured in the high energy / high momentum / small distance limit $x - x = 0$. (Ultraviolet is used by analogy to light, where UV is high energy compared to the visible spectrum; similarly Infrared, or IR, is used to describe long wavelength / low momentum / low energy properties).
To deal with UV divergences, we use a method called... renormalisation! But despite appearing at the same time in these notes, the two uses are essentially unrelated. Let's look at what I'll call Renormalisation II.
First, let's write out the problematic integral explicitly using the expression for the Green's function

$$G_{xy} = \int \mathrm{d}^4p \frac{i}{p^2 - m^2}\exp(-2\pi i p\cdot(x - y)) \tag{533}$$

to give

$$i\frac{\lambda}{2}\int \mathrm{d}^4x G_{1x}G_{xx}G_{x2} = -i\frac{\lambda}{2}\left(\int \mathrm{d}^4p\frac{i}{p^2 - m^2}\right)\int \mathrm{d}^4q\left(q^2 - m^2\right)^{-2}\exp(2\pi i q\cdot(x_1 - x_2)). \tag{534}$$

To get an idea of how these integrals behave, we can use a simple dimensional analysis. Roughly, we expect an integral of the form

$$\int_0^\infty \mathrm{d}^{D+1}p p^n \tag{535}$$

to be convergent for $n < -D - 1$ and divergent for $n > -D - 1$. For $n = -D - 1$ we expect to get a logarithm, which will diverge with infinite limits (log-divergent: the slowest type of divergence).
The first step to dealing with it is *regularisation*. The integrals diverge because of the infinite limts. But in reality, momenta cannot be inifinitely large, and distances cannot be infinitely small. In an

extreme limit, distances cannot be smaller than the Planck length $l_D = \sqrt{G}$ (where G is Newton's constant, in natural units). But we don't need to go to such extremes: our field theory is a coarse-graining, and will only be valid down to some distance / up to some momentum. Call the upper momentum cutoff Λ, which is vastly larger than any momentum we care about but is not infinite (and is much smaller than the ultimate limit set by Planckian units).
Returning to our divergent integral

$$\int \mathrm{d}^4 p \frac{i}{p^2 - m^2} \tag{536}$$

it is convenient to Wick rotate to Euclidean space:

$$(E, \boldsymbol{p}) \to \left(-ip^0, \boldsymbol{p}\right) \tag{537}$$

giving

$$-\int \mathrm{d}^4 \boldsymbol{p} \frac{1}{\boldsymbol{p}^2 + m^2} \tag{538}$$

and use 4D spherical polar co-ordinates to give

$$-2\pi^2 \int_0^\Lambda \mathrm{d}p \frac{p^3}{p^2 + m^2} = -\pi^2 \left[\Lambda^2 - m^2 \ln\left(m^2 + \Lambda^2\right)\right] \tag{539}$$

$$\approx -\pi^2 \left[\Lambda^2 - 2m^2 \ln\left(\Lambda\right)\right] \tag{540}$$

which contains a quadratic divergence and a log divergence.
From here, we simply renormalise (II) by subtracting off this infinite amount. We do this by introducing a 'counterterm' in the Lagrange density:

$$\mathscr{L} = \frac{1}{2}\partial^\mu \varphi \partial_\mu \varphi - \frac{1}{2}\left(m^2 + \delta_m^2\right)\varphi^2 - \frac{\lambda}{4!}\varphi^4 \tag{541}$$

so that when the integral is redone, we find $m^2 \to m^2 + \delta_m^2$ and

$$\int \mathrm{d}^4 p \frac{i}{p^2 - m^2} \to \int \mathrm{d}^4 p \frac{i}{p^2 - m^2 - \delta_m^2} \tag{542}$$

$$= -\pi^2 \left[\Lambda^2 - 2\left(m^2 + \delta_m^2\right)\ln\left(\Lambda\right)\right]. \tag{543}$$

We'd like this to be finite; let's chose it to be $-\pi^2\mu^2$ for some μ. This requires

$$\mu^2 = \Lambda^2 - 2\left(m^2 + \delta_m^2\right)\ln\left(\Lambda\right) \tag{544}$$

or

$$\delta_m^2 = \frac{\mu^2 - \Lambda^2}{-2\ln\left(\Lambda\right)} - m^2. \tag{545}$$

Putting it all together, we find that the corrected Lagrange density in Eq 541 gives the first-order correction to the 2-point function to be

$$i\frac{\lambda}{2}\int \mathrm{d}^4 x G_{1x} G_{xx} G_{x2} = i\frac{\lambda}{2}\pi^2\mu^2 \int \mathrm{d}^4 q \left(q^2 - m^2\right)^{-2} \exp\left(2\pi i q \cdot (x_1 - x_2)\right). \tag{546}$$

By construction, the UV divergence has been subtracted off, by renormalising the mass (through an infinite amount!).
In general, there is a process for identifying μ self-consistently by looking at how it changes the parameters of the theory (specifically the coupling λ). In the particular case above, the 1-loop correction to the 2-point function in φ^4 theory, we can actually choose $\mu = 0$. That is, the propagator does not renormalise at 1st order. The reason is that G_{xx} is special, as it has no momentum. Its Feynman diagram is called a 'tadpole' (when drawn correctly). In general, tadpoles can be renormalised away for free.

If we look at the philosophy of Renormalisation II, we see that it is secretly quite similar to Renormalisation I. The idea is that the bare parameters we started with – in this case m^2 – are not the physical parameters. The things we measure have already been renormalised, and so it is $m^2 + \delta m^2$ which is physical. Still, the mathematical process and reasoning behind Renormalisation I and II are appear rather different.

5 Types of Field and their particles

So far we have focussed on the QFT of a real scalar field φ with quartic interactions. The aim was to go into detail about the calculational techniques of QFT, which are largely the same for all QFTs, without being distracted by nuances of particular field types. In this chapter we will look at some of the features that vary between different field types.

5.1 Complex scalar fields

5.1.1 Recap: complex conjugates in quantum mechanics

Consider the single-particle Schroedinger equation

$$i\frac{\partial\psi}{\partial t} = \hat{H}\psi. \tag{547}$$

Assume the Hamiltonian is real: this is always possible in the absence of a magnetic field. We can take the complex conjugate to give

$$-i\frac{\partial\psi^*}{\partial t} = \hat{H}\psi^* \tag{548}$$

or equivalently

$$i\frac{\partial\psi^*}{\partial(-t)} = \hat{H}\psi^*. \tag{549}$$

This tells us a general property of quantum theories: complex conjugation is equivalent to time reversal. If ψ evolves according to the TDSE, then ψ^* evolves according to the same equation with time reversed. Now consider the current density associated with a particle of charge q

$$\boldsymbol{j}(\boldsymbol{x},t) = \frac{q}{2m}\left\{\psi^*\hat{\boldsymbol{p}}\psi - \psi\hat{\boldsymbol{p}}\psi^*\right\} \tag{550}$$

which obeys the continuity equation

$$\dot{\rho} = -\nabla\cdot\boldsymbol{j} \tag{551}$$

which says that if the charge density ρ changes in a region it must be due to a flow of current into or out of the region. Under complex conjugation $\boldsymbol{j}$ switches sign. Hence, a flow of ψ particles into the region is equivalent to a flow of ψ^* particles out of the region, and vice versa. Equivalently, we can force $\boldsymbol{j}$ to keep the same sign under time reversal by also switching charge $q \to -q$.
This reasoning reveals that ψ^* is the *antiparticle* to ψ: it has the opposite charge but same mass. A flow of particles forwards in time is completely equivalent to a flow of antiparticles backwards in time. Let's see how this works in QFT.

5.1.2 Complex Klein Gordon action

We can define a Klein Gordon theory over complex scalar fields. Now φ and φ^* are treated as independent fields. The action is

$$S^{\mathbb{C}}_{\mathrm{KG}}[\varphi^*,\varphi] = \int \mathrm{d}^4x\left(\partial^\mu\varphi^*\partial_\mu\varphi - m^2\varphi^*\varphi\right). \tag{552}$$

Note that it is conventional to multiply by two relative to Eq 350. Now we can find two Euler Lagrange equations, by taking functional derivatives with respect to our two different fields φ and φ^*. We start by extremising the action with respect to φ^*.
We can do this explicitly, as in Section 3.4.1. But for the sake of variety let's instead use the functional differentiation method.

$$\frac{\delta S^{\mathbb{C}}_{\mathrm{KG}}\left[\varphi^*,\varphi\right]}{\delta\varphi^*\left(y\right)} = \frac{\delta}{\delta\varphi^*\left(y\right)}\int \mathrm{d}^4x \left(\partial^\mu\varphi^*\left(x\right)\partial_\mu\varphi\left(x\right) - m^2\varphi^*\left(x\right)\varphi\left(x\right)\right) \tag{553}$$

$$= \frac{\delta}{\delta\varphi^*\left(y\right)}\int \mathrm{d}^4x \left(-\varphi^*\left(x\right)\partial^2\varphi\left(x\right) - m^2\varphi^*\left(x\right)\varphi\left(x\right)\right) \tag{554}$$

$$= \int \mathrm{d}^4x \left(-\frac{\delta\varphi^*\left(x\right)}{\delta\varphi^*\left(y\right)}\partial^2\varphi\left(x\right) - m^2\frac{\delta\left(\partial^\mu\varphi^*\left(x\right)\right)}{\delta\varphi^*\left(y\right)}\varphi\left(x\right)\right) \tag{555}$$

now use

$$\frac{\delta\varphi\left(x\right)}{\delta\varphi\left(y\right)} = \frac{\delta\varphi^*\left(x\right)}{\delta\varphi^*\left(y\right)} = \delta^4\left(x-y\right) \tag{556}$$

$$\frac{\delta\varphi^*\left(x\right)}{\delta\varphi\left(y\right)} = \frac{\delta\varphi\left(x\right)}{\delta\varphi^*\left(y\right)} = 0 \tag{557}$$

to give

$$\frac{\delta S^{\mathbb{C}}_{\mathrm{KG}}\left[\varphi^*,\varphi\right]}{\delta\varphi^*\left(y\right)} = \int \mathrm{d}^4x\delta^4\left(x-y\right)\left(-\partial^2\varphi\left(x\right) - m^2\varphi\left(x\right)\right) \tag{558}$$

$$= -\partial^2\varphi\left(y\right) - m^2\varphi\left(y\right) \tag{559}$$

where ∂_μ now means $\partial/\partial y^\mu$. Setting both sides equal to zero and relabelling $y \to x$ gives the Klein Gordon equation, as before:

$$\left(\partial^2 + m^2\right)\varphi = 0. \tag{560}$$

Instead taking the functional derivative with respect to φ shows that the complex conjugate field also obeys the Klein Gordon equation:

$$\frac{\delta S^{\mathbb{C}}_{\mathrm{KG}}\left[\varphi^*,\varphi\right]}{\delta\varphi\left(x\right)} = 0 \implies \left(\partial^2 + m^2\right)\varphi^* = 0. \tag{561}$$

The fact that the same mass enters the the equation for the particle and antiparticle is a general feature of QFTs. It is not generally true that the antiparticles obey the same equation of motion as the particles: in general, the equations are complex conjugates.

5.1.3 Complex Klein Gordon Propagator

The propagator for complex Klein Gordon theory is given by

$$G_{xy} = \langle\Omega|\mathcal{T}\hat{\varphi}^\dagger_{\boldsymbol{y}}\hat{\varphi}_{\boldsymbol{x}}|\Omega\rangle. \tag{562}$$

Here, $\hat{\varphi}^\dagger_{\boldsymbol{y}}$ creates a particle, or equivalently annihilates an antiparticle, at y. Similarly, $\hat{\varphi}_{\boldsymbol{x}}$ annihilates a particle/creates an antiparticle at x. In general in QFT we have that

the particle creation operator is an antiparticle annihilation operator, and vice versa.

5.1.4 Microcausality in complex field theories

The previous statement gives a neat interpretation to microcausality (Section) in complex field theories. We now require

$$\langle\Omega|\left[\hat{\varphi}^\dagger_{\boldsymbol{y}}, \hat{\varphi}_{\boldsymbol{x}}\right]|\Omega\rangle = 0 \tag{563}$$

whenever x and y are spacelike separated. Equivalently, we need

$$\langle\Omega|\hat{\varphi}^\dagger_{\boldsymbol{y}}\hat{\varphi}_{\boldsymbol{x}}|\Omega\rangle = \langle\Omega|\hat{\varphi}_{\boldsymbol{x}}\hat{\varphi}^\dagger_{\boldsymbol{y}}|\Omega\rangle. \tag{564}$$

But we see that this requirement is

$$G_{xy}^{\text{particle}} = G_{yx}^{\text{antiparticle}}. \tag{565}$$

Hence, while there is a non-zero amplitude for a particle to propagate across the spacelike region $x-y$, as there was in single-particle QM, in QFT there is always an identical amplitude for the antiparticle to propagate across $y-x$ which perfectly cancels this effect. In this way, antiparticles can be seen as a necessary feature of QFTs which ensure causality.
Note that in the real field theories we considered until now, the particle is its own antiparticle. Therefore, it had to be charge-neutral.

5.1.5 Wick's Theorem in complex field theories

Since φ and φ^* are independent fields, it follows that

$$\langle\Omega|\mathcal{T}\hat{\varphi}_{\boldsymbol{y}}\hat{\varphi}_{\boldsymbol{x}}|\Omega\rangle = 0 \tag{566}$$

$$\langle\Omega|\mathcal{T}\hat{\varphi}^\dagger_{\boldsymbol{y}}\hat{\varphi}^\dagger_{\boldsymbol{x}}|\Omega\rangle = 0. \tag{567}$$

The first line follows because $\hat{\varphi}_{\boldsymbol{x}}|\Omega\rangle = 0$ by definition of the vacuum; the second line follows similarly from $\langle\Omega|\hat{\varphi}^\dagger_{\boldsymbol{y}} = 0$. Combined with Eq 562 for the Green's function, we see that Wick's theorem works the same way in complex field theories, except that the only non-zero contractions involve creation-annihilation pairs.

5.1.6 $U(1)$ charge conservation

The previous equations can be interpreted another way. They say that pairs of particles or antiparticles cannot be created from the vacuum (since the expectation value of a pair of creation or annihilation operators is zero). Similarly we can interpret the Green's function as a finite amplitude to create a particle-antiparticle pair from the vacuum.
Together, this tells us that we have a conservation law: the number of particles minus the number of antiparticles is constant in any process. But recall that an antiparticle has opposite charge to a particle. So this conservation law is simply the conservation of electric charge! Of course, our scalar field is not exactly the field of an electron, so you might ask whether it is really the electric charge we are familiar with. In general it is what we call a $U(1)$ charge, of which electric charge is an example.
To see where the name comes from, note that charge conservation here follows from the fact that only $\varphi^*\varphi$ terms appear in the action, rather than φ^2 or φ^{*2}. Hence, the field can be rescaled by an arbitrary position-dependent complex phase without changing the action:

$$\varphi_x \to \varphi_x \exp\left(i\alpha\left(x\right)\right) \tag{568}$$

$$\varphi^*_x \to \varphi^*_x \exp\left(-i\alpha\left(x\right)\right) \tag{569}$$

$$S^{\mathbb{C}}_{\text{KG}} \to S^{\mathbb{C}}_{\text{KG}} \tag{570}$$

(check this yourself: the kinetic terms seem to change, but only by a total derivative which therefore vanishes under integration). The terms $\exp\left(i\alpha\right)$ for real α describe the set of points on the unit circle; under multiplication this describes the group $U(1)$.

5.1.7 Feynman diagrams and $U(1)$ charge conservation

To keep track of conserved $U(1)$ charge in Feynman diagrams, we add an arrow on the propagator indicating the forwards-in-time propagation of a particle (equivalently, the backwards-in-time propagation of an antiparticle).

Figure 13:

5.1.8 Complex φ^4 theory

We can add interactions as with the real scalar field theory. For example, complex φ^4 theory is described by the interaction term

$$S_{\text{int}} = \frac{\lambda}{4}\int |\varphi_x|^4 \,\mathrm{d}^{D+1}x = \frac{\lambda}{4}\int (\varphi_x^*\varphi_x)^2 \,\mathrm{d}^{D+1}x. \tag{571}$$

Note that the prefactor is $\lambda/4$, not $\lambda/4!$ as it was in the real scalar field theory. The reason is that the symmetry factor of the diagrams is different, owing to the $U(1)$ conservation arrows, or equivalently owing to the fact that we only contract particle-antiparticle pairs in Wick's theorem. For the real field we had

$$\langle \mathcal{T}\varphi_1\varphi_2\varphi_3\varphi_4\rangle_1^{\text{connected}} = i\frac{\lambda}{4!}\int \mathrm{d}^{D+1}x \left\langle \mathcal{T}\varphi_x^4\varphi_1\varphi_2\varphi_3\varphi_4\right\rangle_0 = i\lambda\int \mathrm{d}^{D+1}x G_{1x}G_{2x}G_{3x}G_{4x} \tag{572}$$

where the $1/4!$ prefactor was introduced to deal with the fact that there are 4 ways to pair φ_1 with a φ_x, three ways to pair φ_2, and so on. Now we have

$$\langle \mathcal{T}\varphi_1^\dagger\varphi_2^\dagger\varphi_3\varphi_4\rangle_1^{\mathbb{C},\text{ connected}} = i\frac{\lambda}{4}\int \mathrm{d}^{D+1}x \left\langle \mathcal{T}\varphi_x^{\dagger 2}\varphi_x^2\varphi_1^\dagger\varphi_2^\dagger\varphi_3\varphi_4\right\rangle_0 = i\lambda\int \mathrm{d}^{D+1}x G_{1x}G_{2x}G_{3x}G_{4x}. \tag{573}$$

Here, $\varphi_1^\dagger$ only has a choice of two φ_x, and φ_3 has a choice of two $\varphi_x^\dagger$. The remaining terms are fixed. The corresponding Feynman diagram is shown in Fig 13. With this normalisation it has symmetry factor 1, as required (since all 4 legs are external).

5.2 Bosons and Fermions

Recall the general argument for the possible behaviours of particles under exchange. Define the exchange operator

$$\hat{P}|1,2\rangle \triangleq |2,1\rangle. \tag{574}$$

Since

$$\hat{P}^2|1,2\rangle = |1,2\rangle \tag{575}$$

it follows that

$$\hat{P} = \pm\hat{\mathbb{I}}. \tag{576}$$

Hence, particles are either symmetric or antisymmetric under exchange. We call the former bosons, and the latter fermions.

5.2.1 Bosons

The particles we have considered until now are all bosons, excitations of bosonic fields. To see this, recall that we started from the harmonic oscillator, whose n^{th} excited state we interpreted in second quantized language as n particles. Similarly, we considered it acceptable to raise a given mode in our Fock space as many times as we liked, giving *e.g.* $|24300\ldots\rangle$. Only bosons have the property that multiple particles can exist in the same quantum state (with the same set of quantum numbers). Implicit in this statement is the relation that

bosonic ladder operators obey:

$$\left[\hat{a}_k^\dagger, \hat{a}_l^\dagger\right] = [\hat{a}_k, \hat{a}_l] = 0 \tag{577}$$

$$\left[\hat{a}_k, \hat{a}_l^\dagger\right] = \delta_{kl}. \tag{578}$$

For example, Fock states like $|24300\ldots\rangle$ can only be meaningful if it doesn't matter in which order we applied the creation operators:

$$|243\rangle = \left(\hat{a}^\dagger_{k=1}\right)^2 \left(\hat{a}^\dagger_{k=2}\right)^4 \left(\hat{a}^\dagger_{k=3}\right)^3 |\Omega\rangle = \left(\hat{a}^\dagger_{k=3}\right)^3 \left(\hat{a}^\dagger_{k=1}\right)^2 \left(\hat{a}^\dagger_{k=2}\right)^4 |\Omega\rangle = \textit{etc.} \tag{579}$$

This is the same argument behind quantum numbers in single particle QM. For example, you can't label a ket with both position and momentum, since the operators don't commute so they cannot be simultaneously well defined.

5.2.2 Fermions

For fermions we must instead define a new operator called

the *anticommutator*:

$$\{A, B\} \triangleq AB + BA. \tag{580}$$

Using this, we have that

fermionic ladder operators obey:

$$\left\{\hat{c}^\dagger_k, \hat{c}^\dagger_l\right\} = \{\hat{c}_k, \hat{c}_l\} = 0 \tag{581}$$

$$\left\{\hat{c}_k, \hat{c}^\dagger_l\right\} = \delta_{kl}. \tag{582}$$

Defining

$$\hat{c}^\dagger|\Omega\rangle = |1\rangle \tag{583}$$

it follows immediately that we can only have 0 or 1 particles in a given quantum state, since

$$\left\{\hat{c}^\dagger, \hat{c}^\dagger\right\} = 0 \tag{584}$$

$$\Downarrow$$

$$\hat{c}^\dagger\hat{c}^\dagger = -\hat{c}^\dagger\hat{c}^\dagger \tag{585}$$

and so if we try to define

$$|2\rangle \overset{?}{\triangleq} \hat{c}^\dagger\hat{c}^\dagger|\Omega\rangle \tag{586}$$

we find

$$\hat{c}^\dagger\hat{c}^\dagger|\Omega\rangle = -\hat{c}^\dagger\hat{c}^\dagger|\Omega\rangle \tag{587}$$

$$|2\rangle = -|2\rangle \tag{588}$$

which implies that $|2\rangle \equiv 0$, where this is really 0 and not the vacuum. Hence, trying to raise a fermionic state a second time annihilates it. This is the Pauli exclusion principle.

5.2.3 Grassman variables & integrals [Peskin & Schroeder Sec 9.5]

Recall that when we move from the canonical quantization picture to the path integral picture, we replace non-commuting field operators $\hat{\varphi}_{\boldsymbol{x}}$ with classical commuting field variables φ_x, which assign a real number φ_x at each spacetime position x. Quantum non-commutation then arises from the fractal nature of the field configurations being functionally integrated over.

All of this carries over to fermionic fields, with commutators replaced with anti-commutators. This is quite strange: fermionic fields ψ_x in the path integral picture assign a number to each point in space. But two of these numbers must now anti-commute! This wasn't a problem with bosonic fields, since both real and complex numbers commute.

We call anti-commuting numbers *Grassman variables*. They obey $\theta\eta = -\eta\theta$. An immediate corollary is that $\theta^2 = \theta^{n>1} = 0$. As a result of this, Grassman algebra is actually very simple. For example, any

function can be defined by its Taylor series (as with real functions), but now that Taylor series must terminate at the linear term! For a, $b \in \mathbb{C}$ we have:

$$f(\theta) \equiv a + b\theta \tag{589}$$

where $\equiv$ indicates that this is an exact equality. Of particular interest is the function

$$\exp\left(-\overline{\theta}A\theta\right) = 1 - \overline{\theta}A\theta \tag{590}$$

where $\overline{\theta}$ is the complex conjugate of θ (an independent variable, so that $\theta\overline{\theta} = -\overline{\theta}\theta \neq 0$). Grassman integrals[8] have some freedom in their definition, but the standard choice leads to:

$$\int \mathrm{d}\theta = 0 \tag{591}$$

$$\int \theta \mathrm{d}\theta = 1. \tag{592}$$

As a result, the Gaussian Grassman integral takes a simple form:

$$\int \mathrm{d}\overline{\theta} \int \mathrm{d}\theta \exp\left(-\overline{\theta}A\theta\right) = \int \mathrm{d}\overline{\theta} \int \mathrm{d}\theta \left(1 - \overline{\theta}A\theta\right) \tag{593}$$

$$= \int \mathrm{d}\overline{\theta} \int \mathrm{d}\theta - \int \mathrm{d}\overline{\theta} \int \mathrm{d}\theta\overline{\theta}A\theta \tag{594}$$

$$= \int \mathrm{d}\overline{\theta} \int \mathrm{d}\theta + A \int \overline{\theta}\mathrm{d}\overline{\theta} \int \theta \mathrm{d}\theta \tag{595}$$

$$= A. \tag{596}$$

Similarly, we have the Grassman Gaussian functional integral:

$$\lim_{N\to\infty} \prod_{i=1}^{N} \prod_{j=1}^{N} \int \mathrm{d}\overline{\theta}_i \mathrm{d}\theta_j \exp\left(-\overline{\theta}_i A_{ij} \theta_j\right) = \det(A) \tag{597}$$

perhaps written more familiarly over fields ψ_x as

$$\int \mathscr{D}\overline{\psi}\mathscr{D}\psi \exp\left(-\int \mathrm{d}^{D+1}x \int \mathrm{d}^{D+1}y \overline{\psi}_x \hat{A}_{xy} \psi_y\right) = \det\left(\hat{A}\right). \tag{598}$$

This is proportional to the inverse of the the bosonic Gaussian functional integral

$$\lim_{N\to\infty} \prod_{i=1}^{N} \prod_{j=1}^{N} \int \mathrm{d}\overline{\varphi}_i \mathrm{d}\varphi_j \exp\left(-\overline{\varphi}_i A_{ij} \varphi_j\right) = \frac{(2\pi)^N}{\det(A)}. \tag{599}$$

The Grassman result is so simple that you will often find theories in which determinants play a key role, otherwise unrelated to QFT, rewritten in 'free fermion' form. That is, you rewrite the determinant as a Grassman Gaussian functional integral, and interpret the physical theory as a QFT of non-interaction fermions!

As with bosonic integrals we have

$$\lim_{N\to\infty} \prod_{i=1}^{N} \prod_{j=1}^{N} \int \mathrm{d}\overline{\theta}_i \mathrm{d}\theta_j \theta_k \overline{\theta}_l \exp\left(-\overline{\theta}_i A_{ij} \theta_j\right) = \det(A) A^{-1}_{kl} \tag{600}$$

$$\lim_{N\to\infty} \prod_{i=1}^{N} \prod_{j=1}^{N} \int \mathrm{d}\overline{\theta}_i \mathrm{d}\theta_j \theta_k \overline{\theta}_l \theta_m \overline{\theta}_n \exp\left(-\overline{\theta}_i A_{ij} \theta_j\right) = \det(A) \left(A^{-1}_{kl} A^{-1}_{mn} + A^{-1}_{kn} A^{-1}_{ml}\right) \tag{601}$$

and so on: Wick's theorem applies as before, with the only change being that it is $\det A$ rather than $\det A^{-1}$ which appears as the prefactor.

[8] More frequently called *Berezin integrals,* although this is the only example I know of a Russian getting something named after them when an Englishman invented it first (there are many examples the other way around). The first known reference was David Candlin, 1956, a decade before Berezin.

5.2.4 Quantum statistics of identical non-interacting particles

Using the exchange properties of bosons and fermions, we can immediately derive their respective exchange statistics. Working in the grand canonical ensemble, the grand partition function is

$$\mathscr{Z}(\mu, V, T) = \sum_{n=0}^{\infty} \exp\left(-\beta\left(E_n - N_n\mu\right)\right) \tag{602}$$

where $\beta = 1/T$, N_n is the number of particles in microstate n, and E_n is the energy of that (V is the volume, which will be held fixed). The mean particle number is then given by

$$\langle N\rangle = \beta^{-1}\left(\frac{\partial \ln \mathscr{Z}}{\partial \mu}\right)_{V,T}. \tag{603}$$

For non-interacting identical bosons the energy to add each particle is the same: $\epsilon - \mu$. Hence

$$\mathscr{Z}_{\text{boson}} = \sum_{n=0}^{\infty} \exp\left(-n\beta\left(\epsilon - \mu\right)\right). \tag{604}$$

Recognising the geometric series, we have

$$\mathscr{Z}_{\text{boson}} = \left(1 - \exp\left(-\beta\left(\epsilon - \mu\right)\right)\right)^{-1} \tag{605}$$

and the mean particle number $\langle N\rangle$ is given by

the *Bose Einstein distribution function:*

$$n_B(\mu, \epsilon, T) = \frac{1}{\exp\left(\beta\left(\epsilon - \mu\right)\right) - 1}. \tag{606}$$

These are Bose-Einstein statistics. At low energies $\beta\left(\epsilon - \mu\right) \to 0^+$ this function diverges: the ground state of a set of identical non-interacting bosons has space for an infinite number of particles. At low energies the particles condense into the ground state, simply by their statistics. The result is a Bose Einstein Condensate (BEC). Examples include superfluids (*e.g.* helium-4 cooled below 2.17K: helium-4 is a composite spin-0 boson) and superconductors (in which fermionic electrons bind into bosonic pairs, which can then condense). Some refer to Bose Einstein condensation as a macroscopic quantum phenomenon.

For fermions we can only have 0 or 1 particle, and so

$$\mathscr{Z}_{\text{fermion}} = \sum_{n=0}^{1} \exp\left(-n\beta\left(\epsilon - \mu\right)\right) \tag{607}$$

$$= 1 + \exp\left(-\beta\left(\epsilon - \mu\right)\right) \tag{608}$$

and the mean particle number is given be

the *Fermi Dirac distribution function:*

$$f(\mu, \epsilon, T) = \frac{1}{\exp\left(\beta\left(\epsilon - \mu\right)\right) + 1}. \tag{609}$$

These are Fermi-Dirac statistics. Adding fermions to a system, each quantum state can only contain one particle. At zero temperature, the first particle goes into the lowest energy state, the second particle into the second state, and so on. The result is the 'Fermi sea': a set of filled states at low energy, with a sharp cut-off to a set of empty states at higher energies. At higher temperatures, the cut-off broadens, and some of the highest-energy particles move to higher energy states, leaving the low energy states occupied. This situation describes the electrons in a metal, for example.

5.2.5 Bogoliubov Transformations

Sometimes the most natural way to describe a physical system involves writing down operators that do not conserve particle number. For example, we might find a Hamiltonian like

$$\hat{H}_{\text{boson}} = \epsilon\left(\hat{a}_1^\dagger\hat{a}_1 + \hat{a}_2^\dagger\hat{a}_2\right) + \lambda\left(\hat{a}_1^\dagger\hat{a}_2^\dagger + \hat{a}_1\hat{a}_2\right). \tag{610}$$

Acting on a state of n_i particles of type i $(i = 1, 2)$ gives

$$\hat{H}_{\text{boson}}|n_1, n_2\rangle = \epsilon\,(n_1 + n_2)\,|n_1, n_2\rangle + \lambda\,(|n_1 + 1, n_2 + 1\rangle + |n_1 - 1, n_2 - 1\rangle)\,. \tag{611}$$

Hamiltonians of this form can always be 'diagonalised' by rewriting in terms of new creation and annihilation operators which the commutation relations but which only involve number conserving terms. This is called a Bogoliubov transformation.
First note that we can use the bosonic commutation relations

$$\left[\hat{a}_i, \hat{a}_j\right] = \left[\hat{a}_i^\dagger, \hat{a}_j^\dagger\right] = 0 \tag{612}$$

$$\left[\hat{a}_i, \hat{a}_j^\dagger\right] = \delta_{ij} \tag{613}$$

to rewrite the Hamiltonian as

$$\hat{H}_{\text{boson}} = \frac{\epsilon}{2}\left(\hat{a}_1^\dagger\hat{a}_1 + \hat{a}_1\hat{a}_1^\dagger + \hat{a}_2^\dagger\hat{a}_2 + \hat{a}_2\hat{a}_2^\dagger\right) - \epsilon \tag{614}$$

$$+ \frac{\lambda}{2}\left(\hat{a}_1^\dagger\hat{a}_2^\dagger + \hat{a}_2^\dagger\hat{a}_1^\dagger + \hat{a}_1\hat{a}_2 + \hat{a}_2\hat{a}_1\right) \tag{615}$$

and can then put this into matrix form:

$$2\hat{H}_{\text{boson}} + 2\epsilon = \left(\hat{a}_1^\dagger, \hat{a}_2, \hat{a}_2^\dagger, \hat{a}_1\right)\begin{pmatrix} \epsilon & \lambda & 0 & 0 \\ \lambda & \epsilon & 0 & 0 \\ 0 & 0 & \epsilon & \lambda \\ 0 & 0 & \lambda & \epsilon \end{pmatrix}\begin{pmatrix} \hat{a}_1 \\ \hat{a}_2^\dagger \\ \hat{a}_2 \\ \hat{a}_1^\dagger \end{pmatrix}. \tag{616}$$

Focus on the top 2×2 block:

$$2\hat{H}^{2\times 2}_{\text{boson}} + 2\epsilon = \left(\hat{a}_1^\dagger, \hat{a}_2\right)\begin{pmatrix} \epsilon & \lambda \\ \lambda & \epsilon \end{pmatrix}\begin{pmatrix} \hat{a}_1 \\ \hat{a}_2^\dagger \end{pmatrix}. \tag{617}$$

We can diagonalise this using:

$$\begin{pmatrix} \hat{a}_1 \\ \hat{a}_2^\dagger \end{pmatrix} = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}\begin{pmatrix} \hat{b}_1 \\ \hat{b}_2^\dagger \end{pmatrix}. \tag{618}$$

But we must maintain the bosonic commutation relations. This requires:

$$\left[\hat{a}_1, \hat{a}_1^\dagger\right] = 1 \tag{619}$$

$$\downarrow \tag{620}$$

$$\left[\alpha\hat{b}_1 + \beta\hat{b}_2^\dagger, \alpha^*\hat{b}_1^\dagger + \beta^*\hat{b}_2\right] = 1 \tag{621}$$

$$\left[\alpha\hat{b}_1, \alpha^*\hat{b}_1^\dagger\right] + \left[\beta\hat{b}_2^\dagger, \beta^*\hat{b}_2\right] = 1 \tag{622}$$

$$|\alpha|^2 - |\beta|^2 = 1 \tag{623}$$

and similarly

$$\left[\hat{a}_2, \hat{a}_2^\dagger\right] = 1 \implies |\delta|^2 - |\gamma|^2 = 1. \tag{624}$$

A convenient parameterisation is therefore

$$\begin{pmatrix} \hat{a}_1 \\ \hat{a}_2^\dagger \end{pmatrix} = \begin{pmatrix} \cosh\theta & \sinh\theta \\ \sinh\theta & \cosh\theta \end{pmatrix}\begin{pmatrix} \hat{b}_1 \\ \hat{b}_2^\dagger \end{pmatrix}. \tag{625}$$

Hence we have

$$2\hat{H}^{2\times 2}_{\text{boson}} + 2\epsilon = \left(\hat{b}_1^\dagger, \hat{b}_2\right) \begin{pmatrix} \cosh\theta & \sinh\theta \\ \sinh\theta & \cosh\theta \end{pmatrix} \begin{pmatrix} \epsilon & \lambda \\ \lambda & \epsilon \end{pmatrix} \begin{pmatrix} \cosh\theta & \sinh\theta \\ \sinh\theta & \cosh\theta \end{pmatrix} \begin{pmatrix} \hat{b}_1 \\ \hat{b}_2^\dagger \end{pmatrix} \tag{626}$$

$$= \left(\hat{b}_1^\dagger, \hat{b}_2\right) \begin{pmatrix} \epsilon\cosh 2\theta + \lambda \sinh 2\theta & \epsilon \sinh 2\theta + \lambda \cosh 2\theta \\ \epsilon \sinh 2\theta + \lambda \cosh 2\theta & \epsilon \cosh 2\theta + \lambda \sinh 2\theta \end{pmatrix} \begin{pmatrix} \hat{b}_1 \\ \hat{b}_2^\dagger \end{pmatrix}. \tag{627}$$

Choosing

$$\tanh 2\theta = -\lambda/\epsilon \tag{628}$$

gives

$$\hat{H}^{2\times 2}_{\text{boson}} = \frac{1}{2}\left(\hat{b}_1^\dagger, \hat{b}_2\right) \begin{pmatrix} \sqrt{\epsilon^2 - \lambda^2} & 0 \\ 0 & \sqrt{\epsilon^2 - \lambda^2} \end{pmatrix} \begin{pmatrix} \hat{b}_1 \\ \hat{b}_2^\dagger \end{pmatrix} - \epsilon \tag{629}$$

or

$$\hat{H}^{2\times 2}_{\text{boson}} = \frac{1}{2}\sqrt{\epsilon^2 - \lambda^2}\left(\hat{b}_1^\dagger \hat{b}_1 + \hat{b}_2^\dagger \hat{b}_2\right) - \epsilon + \frac{1}{2}\sqrt{\epsilon^2 - \lambda^2}. \tag{630}$$

Hence, in terms of the new $\hat{b}_i^\dagger$ quasiparticles, the Hamiltonian is diagonal.
We can do the same for Fermions by switching commutators for anti-commutators. Defining

$$\hat{H}_{\text{fermion}} = \epsilon\left(\hat{c}_1^\dagger \hat{c}_1 + \hat{c}_2^\dagger \hat{c}_2\right) + \lambda\left(\hat{c}_1^\dagger \hat{c}_2^\dagger + \hat{c}_2 \hat{c}_1\right) \tag{631}$$

(note the switch of order in the final term, required for Hermiticity) and using the anti-commutation relations

$$\left\{\hat{c}_i, \hat{c}_j\right\} = \left\{\hat{c}_i^\dagger, \hat{c}_j^\dagger\right\} = 0 \tag{632}$$

$$\left\{\hat{c}_i, \hat{c}_j^\dagger\right\} = \delta_{ij} \tag{633}$$

to rewrite as

$$\hat{H}_{\text{fermion}} = \frac{\epsilon}{2}\left(\hat{c}_1^\dagger \hat{c}_1 - \hat{c}_1 \hat{c}_1^\dagger + \hat{c}_2^\dagger \hat{c}_2 - \hat{c}_2 \hat{c}_2^\dagger\right) + \epsilon + \frac{\lambda}{2}\left(\hat{c}_1^\dagger \hat{c}_2^\dagger - \hat{c}_2^\dagger \hat{c}_1^\dagger + \hat{c}_2 \hat{c}_1 - \hat{c}_1 \hat{c}_2\right) \tag{634}$$

and

$$2\hat{H}_{\text{fermion}} - 2\epsilon = \left(\hat{c}_1^\dagger, \hat{c}_2, \hat{c}_2^\dagger, \hat{c}_1\right) \begin{pmatrix} \epsilon & \lambda & 0 & 0 \\ \lambda & -\epsilon & 0 & 0 \\ 0 & 0 & \epsilon & -\lambda \\ 0 & 0 & -\lambda & -\epsilon \end{pmatrix} \begin{pmatrix} \hat{c}_1 \\ \hat{c}_2^\dagger \\ \hat{c}_2 \\ \hat{c}_1^\dagger \end{pmatrix}. \tag{635}$$

Let's again consider the upper 2×2 block (the other now takes a slightly different form):

$$2\hat{H}^{2\times 2}_{\text{fermion}} - 2\epsilon = \left(\hat{c}_1^\dagger, \hat{c}_2\right) \begin{pmatrix} \epsilon & \lambda \\ \lambda & -\epsilon \end{pmatrix} \begin{pmatrix} \hat{c}_1 \\ \hat{c}_2^\dagger \end{pmatrix}. \tag{636}$$

This time it turns out (using the same reasoning as before) that to maintain anticommutation relations requires

$$\begin{pmatrix} \hat{c}_1 \\ \hat{c}_2^\dagger \end{pmatrix} = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} \hat{d}_1 \\ \hat{d}_2^\dagger \end{pmatrix}. \tag{637}$$

This gives

spin	statistics	field	examples
0	boson	scalar	Higgs boson
1/2	fermion	spinor	electron, proton
1	boson	vector	$W^{\pm}$, Z bosons
2	boson	tensor	graviton?

Table 1: Particles labelled by spin. If we succeed in quantizing gravity, we know that the particle excitations will be spin-2.

$$2\hat{H}^{2\times 2}_{\text{fermion}} - 2\epsilon = \cos 2\theta \left(\hat{d}_1^\dagger, \hat{d}_2\right) \begin{pmatrix} \epsilon - \lambda \tan 2\theta & \epsilon \tan 2\theta + \lambda \\ \epsilon \tan 2\theta + \lambda & \lambda \tan 2\theta - \epsilon \end{pmatrix} \begin{pmatrix} \hat{d}_1 \\ \hat{d}_2^\dagger \end{pmatrix} \tag{638}$$

and choosing

$$\tan 2\theta = -\lambda/\epsilon \tag{639}$$

gives

$$2\hat{H}^{2\times 2}_{\text{fermion}} - 2\epsilon = \left(\hat{d}_1^\dagger, \hat{d}_2\right) \begin{pmatrix} \sqrt{\epsilon^2 + \lambda^2} & 0 \\ 0 & -\sqrt{\epsilon^2 + \lambda^2} \end{pmatrix} \begin{pmatrix} \hat{d}_1 \\ \hat{d}_2^\dagger \end{pmatrix}. \tag{640}$$

This Fermionic case occurs in superconductors. Here, the two species of fermion are electrons with opposite crystal momentum. These pair up to form bosonic 'Cooper pairs', created by $\hat{c}_k^\dagger \hat{c}_{-k}^\dagger$, which then Bose-condense in their ground state. There is an energy gap separating the ground state from the first excited state, which is the energy required to break one of these Cooper pairs back into fermions. Interestingly, however, the Bogoliubov transformation shows that the fermions formed by breaking a Cooper pair are *not* electrons! They are superpositions of electrons and holes:

$$\hat{d}_1^\dagger = \cos\theta \hat{c}_k - \sin\theta \hat{c}_{-k}^\dagger. \tag{641}$$

But electrons and holes have opposite charge (the are antiparticles of one another, in a condensed matter setting). Hence these particles, called Bogoliubons, have an electric charge that is in a quantum superposition! When measured, they are found to have any charge from e to $-e$.

5.2.6 The spin-statistics theorem

The exchange operator argument explains that there must be exactly two types of particle in 3D: those that are symmetric or antisymmetric under exchange. We call these bosons and fermions. As we saw in Section X, their exchange properties lead to different statistics: Bose-Einstein for bosons, Fermi-Dirac for fermions.

However, bosons and fermions are also classified by their spin: bosons have integer spin, while fermions have odd-half-integer spin. This perfect correspondence between spin and statistics was for a long time called the spin-statistics connection. In fact it has now been proven, using axiomatic QFT, to be a consequence of CPT symmetry and Lorentz invariance, but the proof is very complicated. The result is the spin-statistics theorem.

Many suspect there should be a simpler explanation for the spin-statistics connection. In a sense, a single spin$-S$ particle has to be rotated through $2\pi/S$ before it returns to its original orientation. Hence, the connection says something about the relationship between the symmetry of a single particle under rotation, versus rotating a pair of particles into one another. So far, no simple explanation of spin-statistics has been identified.

The spin of a particle is deeply connected to the symmetries of spacetime. Specifically, spin labels which irreducible representation of the Poincare group a particle lives in. You will see more of this in the second part of the course. For now, we will proceed to list some basic properties of the different irreps. Table 1 provides a summary.

5.3 A Field Guide to Fields

5.3.1 Spin-0: Scalar fields

Spin-0 fields are Lorentz scalars:

$$\varphi(x) \to \varphi'(x') = \varphi\left(\Lambda^{-1}x\right) = \varphi\left(\Lambda^{\nu}_{\mu}x^{\mu}\right) \tag{642}$$

(*i.e.* the field itself does not transform under the Lorentz transformation). They have been our main focus until now. The non-interacting action is the Klein Gordon action. For a real field this is:

$$\mathscr{L}_{\text{spin}-0} = -\frac{1}{2}\partial^{\mu}\varphi\partial_{\mu}\varphi - \frac{1}{2}m^2\varphi^2. \tag{643}$$

Its Euler Lagrange equation is the Klein Gordon equation, which governs the motion of a non-interacting massive spin-0 particle:

$$\left(\partial^2 + m^2\right)\varphi = 0. \tag{644}$$

The only spin-0 particle in the standard model is the Higgs boson, which obeys the Klein Gordon equation with an added potential (and also interacts with other particles).

5.3.2 Spin-1: Vector fields

Spin-1 fields transform as vectors under Lorentz transformations:

$$A^{\mu}(x) \to A'^{\mu}(x') = \Lambda^{\mu}_{\nu}A^{\nu}\left(\Lambda^{-1}x'\right). \tag{645}$$

The corresponding Lagrange density is

$$\mathscr{L}_{\text{spin-1}} = -\frac{1}{4}F^{\mu\nu}F_{\mu\nu} + \frac{1}{2}m^2A^{\mu}A_{\mu} \tag{646}$$

where we have used the *field strength:*

$$F^{\mu\nu} \triangleq \partial^{\mu}A^{\nu} - \partial^{\nu}A^{\mu}. \tag{647}$$

The corresponding Euler-Lagrange equation is the Proca equation

$$\partial_{\nu}F^{\mu\nu} + m^2A^{\mu} = 0. \tag{648}$$

This describes massive spin-1 particles, such as the W^+, W^-, and Z bosons.

5.3.3 Abelian Gauge fields (spin-1)

With $m = 0$, the Proca equation becomes Maxwell's equations:

$$\mathscr{L}_{\text{Maxwell}} = -\frac{1}{4}F^{\mu\nu}F_{\mu\nu} \tag{649}$$

$$\downarrow \tag{650}$$

$$\partial_{\nu}F^{\mu\nu} = 0. \tag{651}$$

These describe a massless spin-1 particle: the photon. In this case (only) the action features a *gauge symmetry*: the action is invariant under a gauge transformation $A^{\mu}(x) \to A^{\mu}(x) + \partial^{\mu}\lambda(x)$. In electromagnetism this has a simple interpretation. We have

$$A^{\mu} = \begin{pmatrix} \phi \\ \boldsymbol{A} \end{pmatrix}^{\mu} \tag{652}$$

where ϕ is the voltage and $\boldsymbol{A}$ is the magnetic vector potential. That is:

$$\boldsymbol{B} = \nabla \times \boldsymbol{A} \tag{653}$$

$$\boldsymbol{E} = -\nabla V - \dot{\boldsymbol{A}}. \tag{654}$$

Hence the gauge transformation is equivalent to

$$\begin{pmatrix} \phi \\ \boldsymbol{A} \end{pmatrix} \to \begin{pmatrix} \phi + \partial^0 \lambda \\ \boldsymbol{A} + \nabla \lambda \end{pmatrix} \tag{655}$$

which leaves $\boldsymbol{B}$ and $\boldsymbol{E}$ unchanged. The invariance of the action to the gauge transformation encodes the fact that the absolute value of voltage has no meaning: only potential differences are physical. Hence we can shift all voltages in the universe by an amount $\partial^0 \lambda$ and must get the same physical results. Similarly, we can shift the magnetic vector potential by a pure gradient and get the same physical situation.

Maxwell theory, of a massless spin-1 field, is therefore an example of an Abelian gauge theory. Gauge fields are of particular importance in QFT, since it is potentials, rather than forces, which appear in quantum mechanics. For example, the non-relativistic Hamiltonian of a single particle in a magnetic field is

$$\hat{H} = \frac{1}{2m} \left(\hat{p} - i\hat{A} \right)^2 + \hat{V}. \tag{656}$$

This carries over to QFT.

5.3.4 Non-Abelian gauge fields (spin-1)

We can promote any field to a vector of fields simply by adding an index. For example, for a scalar field:

$$\varphi_a(x) = \begin{pmatrix} \varphi_1(x) \\ \varphi_2(x) \\ \vdots \\ \varphi_N(x) \end{pmatrix}. \tag{657}$$

This index is not a spacetime index, so this does not describe a spin-1 field. It is simply a convenient notation for handling multiple independent fields. You can, for example, rewrite a single complex scalar field as a vector of real fields with an appropriate metric. We can just as easily add a vector index to a spin-1 field, to describe N independent spin-1 fields:

$$A_a^\mu(x) = \begin{pmatrix} A_1^\mu(x) \\ A_2^\mu(x) \\ \vdots \\ A_N^\mu(x) \end{pmatrix}. \tag{658}$$

Such fields appear in the Standard Model. Here, the vector space indexed by a is the non-Abelian gauge group $G = SU(2)$ or $SU(3)$; then $N = \dim G$. The result is described by Yang Mills theory:

$$\mathscr{L}_{\text{Yang-Mills}} = -\frac{1}{4} F_a^{\mu\nu} F_{\mu\nu}^a \tag{659}$$

where

$$F_a^{\mu\nu} \triangleq \partial^\mu A_a^\nu - \partial^\nu A_a^\mu + g f^{abc} A_b^\mu A_c^\nu. \tag{660}$$

The corresponding Euler-Lagrange equation is

$$D_\nu F^{\mu\nu} = 0 \tag{661}$$

where

$$D^\mu \triangleq \partial^\mu - igT^a A_a^\mu. \tag{662}$$

Here T^a are the generators, and f^{abc} the structure constants, of the Lie algebra of G:

$$\left[T^a, T^b \right] = i f^{abc} T^c \tag{663}$$

(since these are not spacetime indices they can be raised and lowered without any change). For the simplest case of $G = SU(2)$, we can choose as generators the Pauli matrices σ. Then we have the familiar relation

$$\left[\sigma^a, \sigma^b\right] = 2i\epsilon^{abc}\sigma^c \tag{664}$$

where ϵ^{abc} is antisymmetric in any pair of indices. Non-Abelian gauge fields describe quarks and gluons in the Standard Model: called quantum chromodynamics, the gauge group is $SU(3)$. Additionally, electroweak theory describing W and Z particles and the photon starts off as an $SU(2)$ non-Abelian gauge theory, although the story here is complicated by the Higgs mechanism.

5.3.5 Spin-1/2

The spin-1/2 field is called a spinor field. The Lagrange density is

$$\mathscr{L}_{\text{spin-1/2}}\left(\overline{\psi}, \psi\right) = \overline{\psi}_a \left(i\gamma^\mu_{ab}\partial_\mu - m\mathbb{I}_{ab}\right)\psi_b \tag{665}$$

where

$$\overline{\psi}_a \triangleq \psi^\dagger_{ab}\gamma^0_b \tag{666}$$

and the corresponding Euler Lagrange equation is the Dirac equation

$$\left(i\gamma^\mu_{ab}\partial_\mu - m\mathbb{I}_{ab}\right)\psi_b = 0. \tag{667}$$

Here, the *gamma matrices* are any matrices obeying the anti-commutation relations:

$$\{\gamma^\mu, \gamma^\nu\} = 2\eta^{\mu\nu}\mathbb{I} \tag{668}$$

which defines the *Clifford algebra* $\mathrm{Cl}_{1,3}(\mathbb{R})$. The smallest dimension of representation is 4, giving 4×4 matrices. In fact any 4×4 representation of this algebra is unitarily equivalent to any other:

$$\gamma^{\mu'} = S^\dagger\gamma^\mu S \tag{669}$$

for some unitary 4×4 matrix S. One choice, the Weyl representation, is

$$\gamma^0 = \sigma^3 \otimes \mathbb{I}_2 \tag{670}$$
$$\gamma^j = i\sigma^2 \otimes \sigma^j. \tag{671}$$

Dirac derived his equation by taking the 'square root' of the Klein Gordon equation, and working out a set of matrices which allowed this to work. Hence, you can square the Dirac equation to return to the Klein Gordon equation:

$$\left(i\gamma^\mu_{ab}\partial_\mu - m\mathbb{I}_{ab}\right)\psi_b = 0 \tag{672}$$
$$\downarrow \tag{673}$$
$$\left(-i\gamma^\nu_{ca}\partial_\nu - m\mathbb{I}_{ca}\right)\left(i\gamma^\mu_{ab}\partial_\mu - m\mathbb{I}_{ab}\right)\psi_b = 0 \tag{674}$$
$$\left(\gamma^\nu_{ca}\gamma^\mu_{ab}\partial_\nu\partial_\mu + m^2\mathbb{I}_{cb}\right)\psi_b = 0 \tag{675}$$
$$\left(\frac{1}{2}\left\{\gamma^\nu_{ca}, \gamma^\mu_{ab}\right\}\partial_\nu\partial_\mu + m^2\mathbb{I}_{cb}\right)\psi_b = 0 \tag{676}$$
$$\left(\partial^\mu\partial_\mu + m^2\right)\psi_c = 0 \tag{677}$$

where it should be noted that ψ_a has the basic appearance of a 4-component vector. It is not a Lorentz vector; rather it is a vector in 'spinor space', although this is linked to Minkowski space. As a result the transformation of spinor fields is a bit more complicated:

$$\psi_a(x) \to \psi'_a(x') = S_{ab}(\Lambda)\,\psi_b\left(\Lambda^{-1}x\right). \tag{678}$$

It is conventional to drop the spinor indices, and to additionally define the 'Feynman scratch notation'

$$\not{a} \triangleq \gamma^\mu a_\mu \tag{679}$$

to write the Dirac equation in the neater form

$$\left(i\not{\partial} - m\right)\psi = 0. \tag{680}$$

5.4 The Standard Model

The Lagrange density for the Standard Model is

$$\mathscr{L}_{\text{SM}} = \mathscr{L}_{\text{boson}} + \mathscr{L}_{\text{fermion}} + \mathscr{L}_{\text{Higgs}} + \mathscr{L}_{\text{Yukawa}} \tag{681}$$

with

$$\mathscr{L}_{\text{boson}} = -\frac{1}{4}G^{a\mu\nu}G^a_{\mu\nu} - \frac{1}{4}W^{a\mu\nu}W^a_{\mu\nu} - \frac{1}{4}B^{\mu\nu}B_{\mu\nu} \tag{682}$$

$$\mathscr{L}_{\text{fermion}} = \sum_{\psi\in\text{fermions}} \overline{\psi} i \not{D}_\mu \psi \tag{683}$$

$$\mathscr{L}_{\text{Higgs}} = \left(D_\mu H\right)^\dagger D^\mu H + \mu^2 H^\dagger H - \lambda\left(H^\dagger H\right)^2 \qquad (\mu^2 > 0). \tag{684}$$

Here,

- $G^a_{\mu\nu}$ is the field strength of the non-abelian $SU(3)$ group (strong force-carrying bosons: gluons)
- $W^a_{\mu\nu}$ is the field strength of the non-abelian $SU(2)$ group (electroweak force-carrying bosons)
- $B_{\mu\nu}$ is the field strength of the abelian $U(1)$ field (hypercharge)
- ψ includes all fermions: quarks and leptons, left- and right-handed
- $\mathscr{L}_{\text{Yukawa}}$ describes the Fermion masses.

This form is before the Higgs field has induced spontaneous symmetry breaking. You now have all the tools to do calculations with these fields and their particles, in the second part of this course.