An Extremely Brief Introduction to Quantum Field Theory

W. A. Horowitz
Department of Physics, University of Cape Town, Private Bag X3, Rondebosch 7701, South Africa
E-mail: wa.horowitz@uct.ac.za

Abstract. We provide a very brief introduction to $n$-dimensional scalar field theory, with an eye to renormalization and expectation values of operators. We assume the audience already has some experience with QFT.

1. Introduction: Making Quantum Mechanics Relativistic

Recall the postulates of non-relativistic quantum mechanics (NRQM):

(i) the state of the system is represented by a vector $|\psi\rangle \in \mathcal{H}$, a Hilbert space
(ii) observables are represented by (essentially) self-adjoint operators that obey the Dirac quantization condition, $\{A, B\}_\text{classical} = \mathcal{C} \rightarrow [\hat{A}, \hat{B}]_\text{commutator} = i\mathcal{C}$
(iii) if the system is in a state $|\psi\rangle$, then the measurement of an observable represented by $\hat{O}$ will be an eigenvalue $\omega$ of $\hat{O}$, $\hat{O}|\omega\rangle = \omega|\omega\rangle$ with probability $P(\omega) \propto |\langle \omega | \psi \rangle|^2$; as a result of the measurement, the state of the system changes instantaneously from $|\psi\rangle$ to $|\omega\rangle$
(iv) $|\psi(t)\rangle$ evolves according to the Schrödinger equation, $i\hbar \frac{\partial}{\partial t} |\psi(t)\rangle = \hat{H}(t)|\psi(t)\rangle$, where $\hat{H}$ is the Hamiltonian operator for the system, which may explicitly (or implicitly) depend on time $t$; we usually take $\hat{H} = \hat{p}^2/2m + V(\hat{x})$.

Now recall the postulates of special relativity (SR):

(i) the laws of physics are the same in all inertial reference frames (inertial frames move with constant velocity with respect to each other; observers in different frames must agree on the results of physical measurements; and the laws of physics must be the same, which is to say that all inertial observers agree on the initial conditions and equations of motion for the system)
(ii) the speed of light is a constant $c = 1$ in all inertial frames, which is equivalent to requiring that our physical models are on a pseudo-Riemannian manifold with metric $\eta^{\mu\nu}$ with diagonal elements $\{+1, -1, -1, -1, \ldots\}$, which is again equivalent to requiring that space and time are treated on equal footing: physical quantities transform under the Lorentz group, for example $U(\Lambda) \phi(x) = \phi(\Lambda^{-1}x)$ for a scalar field; finally, the speed of light acts as a speed limit—information travels (at most) as fast as the speed of light in vacuum.

Notice that some of the postulates of NRQM appear to conflict with the postulates of SR. In particular, wavefunction collapse in NRQM is immediate and the Schrödinger equation is
not relativistically invariant. The issue of instantaneous wavefunction collapse was first raised by Einstein, Podolsky, and Rosen [1]. We will avoid the issue in two ways: first, wavefunction collapse transmits no information (and therefore does not violate the speed limit imposed by the finite speed of light); and second, in quantum field theory (QFT), we only ask questions about asymptotic past or future states of the system.

The Schrödinger equation breaks Lorentz invariance in two ways. First, the usual Hamiltonian
\[ \hat{H} = \frac{\hat{p}^2}{2m} + V(\hat{x}) \]
does not yield the relativistically correct dispersion relation. One may readily modify the Hamiltonian to \( \hat{H} = \sqrt{\hat{p}^2 + m^2} \), leading to the correct relation \( E = \sqrt{p^2 + m^2} \).

However, the Schrödinger equation breaks Lorentz invariance in a second, fundamental way by treating time and space as fundamentally different: the Schrödinger equation is linear in \( \partial_t \) but quadratic in \( \partial_x \); similarly, position is elevated to an operator while time is only a parameter.

In order to resolve the conflict between NRQM and SR we have two options: promote \( t \) to an operator (the string theory like approach [2,3]) or denote \( \hat{x} \) to a parameter and treat fields as operators (particle theory like approach [4–12]). As this is a particle physics context, we will follow the latter approach.

2. The Relativistic Quantum Scalar Field

We will take symmetries as the fundamental organizing principle for QFT.

In path integral language, we seek a Lagrangian that respects the symmetries we postulate for our physics. (Path integrals perfectly preserve classically the symmetries of the Lagrangian explicitly. One finds that quantum mechanics leads to anomalies: renormalization can break classical symmetries.)

Let’s start by postulating that scalar particles are the quanta of a field \( \phi \) whose Lagrangian is
\[ \mathcal{L} = \frac{1}{2} (\partial_{\mu} \phi_0)^2 - \frac{1}{2} m_0^2 \phi_0^2 + \lambda_0 \phi_0^4. \] (1)

Scalars are valuable because they are the simplest mathematical objects to work with (c.f. fermions, vector fields, etc.). Scalars also describe a wealth of physics (e.g. the Higg’s boson and scalar and pseudoscalar mesons).

Eq. (1) is written in terms of bare quantities \( \phi_0 \), \( m_0 \), and \( \lambda_0 \) in anticipation of performing renormalized perturbation theory, in which we will re-write the Lagrangian in terms of renormalized quantities \( \phi_r \), \( m_r \), and \( \lambda_r \).

Renormalization is a generic feature of QFT (not a bug!). Mathematically, there are infinities that need to be rigorously dealt with. Physically, we require that the parameters in our theory are connected in some way to a physical measurement. Additionally, our theories will only make sense if the physics at one scale is not affected by the physics at a dramatically different scale. (For example, the physics at infinitesimally small distances shouldn’t alter the dynamics at macroscopic lengths.)

Remember that we are dealing with an inherently quantum mechanical system, which implies Heisenberg’s Uncertainty Principle, which implies that virtual pair production is possible; i.e. particle–anti-particle pairs pop in and out of existence all the time. Our theory is self-interacting. So for any one physical particle \( \rho \) of our theory, there is a cloud of virtual particles...
surrounding it, all interacting with each other and with $p$; see Fig. [1]. In fact, physically, we cannot tell the difference between $p$ and its cloud. So what an experimentalist measures for the charge of $p$ (related to $\lambda_0$) or the mass of $p$ (related to $m_0$) or even how many $p$’s there are (related to $\phi_0$) may depend on the energy scale (or, equivalently, distance) at which they probe $p$.

### 3. Quantizing the Free Relativistic Scalar Field

Our goal is to quantize the free relativistic scalar field in $n$ spatial dimensions. (Note that we will use a more naturally relativistic notation than Peskin and Schroeder [7], although we will keep the particle physicist’s mostly minus convention for the metric.)

We will postulate that the relevant relativistically invariant Lagrange (density) is

$$\mathcal{L} = \frac{1}{2} (\partial_{\mu} \hat{\phi}_0)^2 - \frac{1}{2} m_0^{2} \hat{\phi}_0^2. \quad (2)$$

We seek a solution $\hat{\phi}_0$ associated with the above Lagrange (density) and a spectrum of the field operator (eigenvalues and eigenvectors). We will find that $\hat{\phi}_0$ is the bare (unrenormalized) operator associated with the production of a single scalar particle of mass $m_0$. For notational simplicity, for the rest of this section we will drop the 0 subscript and the operator hat. (Note that since the theory is non-interacting, we can solve the theory exactly and without renormalization.)

Dimensional analysis is an invaluable tool for the physicist, and so we should pause for a moment to analyze the dimensions of the objects in our theory. The path integral goes as $\int D\phi e^{iS}$, where $S = \int d^{n+1}\mathbf{L}$. Since one can only take a number to the power of a dimensionless number, the action must be dimensionless, $|S| = 1$. Since $[d^{n+1}x] = L^{n+1} = E^{-(n+1)}$, we have that $[\mathcal{L}] = E^{n+1}$. Since $[\partial_{\mu}] = [m] = E$, we have that $[\phi] = E^{(n-1)/2}$.

Our solution $\phi$ must first of all satisfy the classical equations of motion found by extremizing the action. The Euler-Lagrange equations yield

$$(\Box + m^2)\phi = 0. \quad (3)$$

Already things are looking very promising: the equations of motion are Lorentz invariant (as must be as the Lagrangian is Lorentz invariant).

We may readily solve Eq. (3) by decomposing our solution into Fourier modes:

$$\hat{\phi}(x) = \int \frac{d^np}{(2\pi)^n2E_p} \left( a_p e^{-ipx} + a^\dagger_p e^{ipx} \right) \gamma^{01} = E_p = \sqrt{p^2 + m^2}. \quad (4)$$

Notice how: 1) we have (explicitly) separated out the classical from the quantum in Eq. (4), where $\hat{\phi}^\dagger = \hat{\phi}$ automatically; 2) the Fourier modes obey the usual relativistic dispersion relation; and 3) dimensional analysis implies that $E^{(n-1)/2} = E^{n-1}[a]$, so then $[a] = E^{(1-n)/2} = [a]$. In order to quantize Eq. (4) we must impose the Dirac quantization condition. We must of course then first decide what the Dirac quantization condition is. Since we are interested in relativistic theories, we will require that fields cannot influence each other outside of the lightcone. Hence a sensible generalization from the 1D NRQM case of $[\hat{x}, \hat{p}] = i$ for the Dirac quantization condition for fields is to require an equal-time contact interaction:

$$[\hat{\phi}(x^0, \vec{x}), \hat{\pi}(x^0, \vec{y})] = i \delta^{(n)}(\vec{x} - \vec{y}); \quad (5)$$

i.e. fields at the same time (in one inertial frame) may only affect each other at exactly equal points (otherwise information could propagate faster than the speed of light).

As an aside, it’s an interesting exercise, once we’ve solved for $\phi(x)$, to compute the commutator $[\hat{\phi}(x), \mathcal{O}(y)]$ without the equal time restriction.
In order to impose our proposed commutator relation Eq. 8 we need to compute the canonically conjugate momentum to the field: \( \pi(x) \equiv \partial L / \partial (\partial_0 \phi) = \partial^0 \phi = \partial \phi \). One may easily see that

\[
\pi(x) = -\frac{i}{2} \int \frac{d^np}{(2\pi)^n} \left( \hat{a}_\rho e^{-ip \cdot x} - \hat{a}^\dagger_\rho e^{ip \cdot x} \right) \big| p^\rho = E_p = \sqrt{p^2 + m^2}. \tag{6}
\]

One may then compute the equal-time commutator

\[
\left[ \hat{\phi}(x^0, \vec{x}) , \pi(x^0, \vec{y}) \right] = \int \frac{d^n p d^n q}{(2\pi)^n (2\pi)^n} \left\{ -[\hat{a}_\rho, \hat{a}^\dagger_\sigma] e^{ig \cdot (y-x)} + [\hat{a}^\dagger_\rho, \hat{a}_\sigma] e^{ig \cdot (x-y)} \right\}. \tag{7}
\]

We postulated that Eq. 7 should equal \( i\delta^{(n)}(\vec{x} - \vec{y}) \), Eq. 5. One may show that if

\[
[\hat{a}_\rho, \hat{a}^\dagger_\sigma] = (2\pi)^n 2E_p \delta^{(n)}(\vec{x} - \vec{y}), \tag{8}
\]

then Eq. 7 does satisfy Eq. 5. Notice that the required commutation relation for the raising and lowering operators Eq. 8 is consistent with the dimensional analysis requirement that \( [\hat{a}_\rho] = [\hat{a}^\dagger_\rho] = E^{(1-n)/2} \).

Now, from a direct application of Noether’s theorem,

\[
T^{\mu \nu} = \frac{\partial L}{\partial (\partial_\mu \phi)} \partial^\nu \phi - \eta^{\mu \nu} L
= \partial^\mu \phi \partial_\nu \phi - \eta^{\mu \nu} \left[ \frac{1}{2} \partial_\alpha \phi \partial^\alpha \phi - \frac{1}{2} m^2 \phi^2 \right]. \tag{9}
\]

Notice that the above \( T^{\mu \nu} \) has the right dimensions for an energy momentum tensor density. In particular, \( E = \int d^nx T^{00} \). Since \( \int d^nx x = E^{-n} \), we have that \( [T^{\mu \nu}] = E^{n+1} \), which is consistent with, e.g., \( \partial_\mu \phi \partial^\nu \phi = [\partial^\mu \phi][\phi] [\partial_\nu \phi] = E E^{(n-1)/2} E E^{(n-1)/2} = E^{n+1} = [m^2 \phi^2] \).

It’s worth noting, however, that Eq. 9 is sometimes not the best energy momentum tensor density to use. For example, Eq. 9 has non-zero trace even for a conformal (massless) scalar field. An improved energy momentum tensor density [6,13] is

\[
\rho^{\mu \nu} \equiv T^{\mu \nu} - \frac{1}{4} \frac{n}{n-1} (\partial^\mu \partial_\nu - \eta^{\mu \nu} \Box) \phi^2. \tag{10}
\]

One may work out from the solution for \( \phi \) and the commutation relations Eq. 8 that

\[
\hat{\rho}^{\mu} = \int d^n x \hat{T}^{\mu \nu} = \int \frac{d^n p}{(2\pi)^n 2p^0} p^\mu \left( \hat{a}^\dagger_\rho \hat{a}_\rho + 2p_0 (2\pi)^n \delta^{(n)}(0) \right) \to \int \frac{d^n p}{(2\pi)^n 2p^0} p^\mu \hat{a}^\dagger_\rho \hat{a}_\rho, \tag{11}
\]

where in the final step we dropped the infinite vacuum energy contribution (the potentially infinite spatial infinities are 0 by symmetry). Remember that in QFT we only care about energy differences, not the absolute energy.

We immediately see that we’ve fully solved our problem: our states are specified by \( | \ldots n_{\hat{p}}, n_{\hat{q}}, \ldots \rangle \propto (\hat{a}^\dagger_{\hat{p}})^{n_{\hat{p}}} (\hat{a}^\dagger_{\hat{q}})^{n_{\hat{q}}} \ldots | 0 \rangle \), where \( | 0 \rangle \) is the ground state (i.e. the vacuum) with \( \hat{a}_\rho | 0 \rangle = 0 \), and each state has energy \( \hat{H} | \ldots n_{\hat{p}}, n_{\hat{q}}, \ldots \rangle = (\ldots + n_{\hat{p}} p^0 + n_{\hat{q}} p^0 + \ldots) | \ldots n_{\hat{p}}, n_{\hat{q}}, \ldots \rangle \).

We may now interpret our result: the quanta that are created and destroyed by \( \hat{a}^\dagger_\rho \) and \( \hat{a}_\rho \) have all the properties we expect of particles (energy, momentum, spin [not shown], come in discrete numbers), so we’ll make the intellectual leap to interpret these quanta as particles.

In particular, let’s examine

\[
\phi(x) | 0 \rangle = \int \frac{d^n p}{(2\pi)^n 2p^0} \left( \hat{a}_\rho e^{-ip \cdot x} + \hat{a}^\dagger_\rho e^{ip \cdot x} \right) | 0 \rangle = \int \frac{d^n p}{(2\pi)^n 2p^0} e^{ip \cdot x} | \vec{p} \rangle, \tag{12}
\]
where we’ve defined $|\vec{p}\rangle \equiv \hat{a}^\dagger_{\vec{p}}|0\rangle$ as the usual momentum eigenstate. (More on $|\vec{p}\rangle$ in a moment.)

Eq. 12 should look very familiar. Recall that in NRQM, $|\vec{x}\rangle = \int \frac{d^nq}{(2\pi)^n} e^{-i\vec{p} \cdot \vec{x}} |\vec{p}\rangle$. Eq. 12 is clearly the relativistic analog: the integration measure is changed to be relativistically invariant and the $E_{\vec{p}}$ giving the time dependence (in the Heisenberg picture) is now relativistically correct.

We will therefore interpret $\phi(x)|0\rangle$ as creating a single particle at position $x^\mu$. Looking towards the future, when we have an interacting theory, we expect the only change to be that $\phi_0(x)$ may not create exactly one particle at $x^\mu$ (or, equivalently, may not create a particle with probability 1); rather, we may need to renormalize the weight of the wavefunctions.

Notice that our momentum eigenstates contain a funny dimension. We’ll take the vacuum to be properly normalized, $\langle 0|0 \rangle = 1$; i.e. $|0\rangle = 1$, the vacuum bra and ket are dimensionless. Since $[\hat{a}^\dagger_{\vec{p}}] = E^{(1-n)/2}$, we have that $[|\vec{p}\rangle] = E^{(1-n)/2}$. Note also that the inner product of two momentum eigenstates is $\langle \vec{k} | \vec{p} \rangle = \langle 0 | \hat{a}^\dagger_{\vec{k}} \hat{a}_{\vec{p}} | 0 \rangle = (2\pi)^n 2E_{\vec{p}} \delta^{(n)}(\vec{k} - \vec{p})$. Thus $[|\vec{k}\rangle | \vec{p} \rangle] = E^{1-n} \Rightarrow [|\vec{p}\rangle] = E^{(1-n)/2}$, consistent with the above.

Now, if $|\vec{p}\rangle$ is really a momentum eigenstate, we should find that

$$\langle \vec{p} | \hat{P}_\mu | \vec{p} \rangle = \frac{\langle \psi | \hat{P}_\mu | \psi \rangle}{\langle \psi | \psi \rangle} |_{|\psi \rangle = |\vec{p}\rangle} = \frac{\langle \vec{p} | \hat{P}_\mu | \vec{p} \rangle}{\langle \vec{p} | \vec{p} \rangle} = p^\mu. \quad (13)$$

Let’s check:

$$\langle \vec{p} | \hat{P}_\mu | \vec{p} \rangle = \frac{\langle 0 | \hat{a}_{\vec{p}} \int d^n x \hat{T}^{\mu 0} \hat{a}_{\vec{p}}^\dagger |0\rangle}{\langle \vec{p} | \vec{p} \rangle}$$

$$= \frac{1}{\langle \vec{p} | \vec{p} \rangle} \langle 0 | \hat{a}_{\vec{p}} \int (2\pi)^n (2q^0)^n q^\mu \hat{a}_{\vec{q}}^\dagger \hat{a}_{\vec{q}} |0\rangle$$

$$= \frac{1}{\langle \vec{p} | \vec{p} \rangle} \int (2\pi)^n 2q^0 q^\mu (0) \hat{\vec{p}} \hat{\vec{q}} \delta^{(n)}(\vec{p} - \vec{q}) |0\rangle$$

$$= p^\mu \langle \vec{p} | \vec{p} \rangle / \langle \vec{p} | \vec{p} \rangle = p^\mu, \quad (14)$$

where to get from the second line to the third line we used Eq. 11.

Since $\langle \vec{p} | \vec{p} \rangle = (2\pi)^n 2E_{\vec{p}} \delta^{(n)}(\vec{p})$, in the last line of the above derivation we formally divided $\infty$ by $\infty$ in order to obtain a finite result. This formal division of an infinite quantity by exactly the same infinite quantity to obtain a finite result is a common procedure in QFT.

Another way in which we could’ve arrived at a finite result above is to recognize that $|\vec{p}\rangle$ is an unphysical state: we haven’t taken into account the quantum fuzziness (Heisenberg uncertainty) in the natural world. Physically, one would at best have access to a normalizable wavepacket of some width. We should really use a state

$$|\psi\rangle = \int \frac{d^n p}{(2\pi)^n 2p^0} f(\vec{p}) |\vec{p}\rangle, \quad (15)$$

where $f(\vec{p})$ is a smooth function whose normalization is given by

$$\langle \psi | \psi \rangle = \int \frac{d^n p d^n q}{(2\pi)^n 2p^0 (2\pi)^n 2q^0} f^\ast(\vec{p}) f(\vec{q}) = \int \frac{d^n p}{(2\pi)^n 2p^0} |f(\vec{p})|^2 = 1. \quad (16)$$

Two related advantages to using Eq. 15 are that: 1) $|\psi\rangle$ is dimensionless, as we are used to; and 2) we don’t need to divide by $\langle \psi | \psi \rangle$ (or, rather, dividing by $\langle \psi | \psi \rangle$ is trivial as it’s 1).
With our $|\psi\rangle$ in hand, we may evaluate

$$\langle \hat{P}^\mu \rangle_{|\psi\rangle} = \langle \psi | \hat{P}^\mu | \psi \rangle = \int \frac{d^n p}{(2\pi)^n 2 p^0} \frac{d^n q}{(2\pi)^n 2 q^0} d^n s f^*(\vec{p}) f(\vec{s}) \langle 0 | \hat{a}_p \hat{a}_q \hat{a}_s | 0 \rangle$$

$$= \int \frac{d^n p}{(2\pi)^n 2 p^0} |f(\vec{p})|^2 p^\mu$$

Now Eq. 16 tells us that $\frac{1}{(2\pi)^n 2 p^0} |f(\vec{p})|^2$ is a probability distribution for the momentum of the particle. Thus Eq. 17 is precisely the average momentum of the particle, which is then the average value of the momentum in the full field. Thus $\langle \hat{P}^\mu \rangle = \langle p^\mu \rangle$. For a momentum distribution whose average value is some momentum $p^\mu_0$ we have $\langle \hat{P}^\mu \rangle = p^\mu_0$. In a similar vein, we may make a very Peskin-like argument that, should $f(\vec{p})$ be highly peaked (i.e. we are experimentally examining an ensemble average of particles of a nearly definite momentum $p^\mu_0$), like at the LHC, then we may replace $\int \frac{d^n p}{(2\pi)^n 2 p^0} |f(\vec{p})|^2 p^\mu \rightarrow p^\mu_0$. Then $\langle \hat{P}^\mu \rangle = p^\mu_0$.

4. Acknowledgments
The author wishes to thank the National Research Foundation and the SA-CERN collaboration for generous support.

References
[1] A. Einstein, B. Podolsky, and N. Rosen, *Phys. Rev.*, vol. 47, pp. 777–780, 1935.
[2] B. Zwiebach, *A first course in string theory*. Cambridge University Press, 2006.
[3] J. Polchinski, *String theory. Vol. 1: An introduction to the bosonic string*. Cambridge University Press, 2007.
[4] C. Itzykson and J. B. Zuber, *Quantum Field Theory*. International Series In Pure and Applied Physics, New York: McGraw-Hill, 1980.
[5] G. F. Sterman, *An Introduction to quantum field theory*. Cambridge University Press, 1993.
[6] L. S. Brown, *Quantum Field Theory*. 1994.
[7] M. E. Peskin and D. V. Schroeder, *An Introduction to quantum field theory*. 1995.
[8] S. Weinberg, *The Quantum theory of fields. Vol. 1: Foundations*. Cambridge University Press, 2005.
[9] A. Zee, *Quantum field theory in a nutshell*. 2003.
[10] M. Srednicki, *Quantum field theory*. Cambridge University Press, 2007.
[11] S. Coleman, “Notes from Sidney Coleman’s Physics 253a: Quantum Field Theory,” 2011.
[12] M. D. Schwartz, *Quantum Field Theory and the Standard Model*. Cambridge University Press, 2014.
[13] C. G. Callan, Jr., S. R. Coleman, and R. Jackiw, “A New improved energy - momentum tensor,” *Annals Phys.*, vol. 59, pp. 42–73, 1970.