

Lecture 1: Perturbative renormalization

Edward Witten¹

October 1996

This lecture is a very elementary introduction to renormalization of Feynman integrals.

1.1. Perturbative expansion of a 2-point correlation function.

Consider a quantum field theory with one scalar bosonic field $\phi(x)$ on a Minkowski space $(V, (\cdot, \cdot))$ of signature $(1, n-1)$ (see Kazhdan's lectures). Let V_+, V_- be the closures of the upper and the lower part of the full cone of time-like vectors. Let \mathcal{H} be the Hilbert space of the theory, $\mathcal{D} \subset \mathcal{H}$ a dense subspace, and $\varphi : \mathcal{S}(V) \rightarrow \text{End}(\mathcal{D})$ be the quantization map. We assume that the triple $(\mathcal{H}, \mathcal{D}, \phi)$ satisfies Wightman axioms. In this case we can define quantum fields $\phi(x) := \varphi(\delta_x)$, which are distributions on V with values in $\text{End}(\mathcal{D})$. When it does not cause confusion, we will treat them as usual $\text{End}(\mathcal{D})$ -valued functions.

We can assume without loss of generality that the 1-point Wightman function of the theory vanishes. Indeed, the 1-point function is a constant C , and we can redefine the map φ by setting $\varphi'(f) = \varphi(f) - C \int f dv$. The map φ' satisfies Wightman axioms as well and has a zero 1-point function.

We look at the time ordered 2-point Wightman function

$$1.1 \quad \mathcal{W}_2^T(x, y) = \langle \Omega | T(\phi(x)\phi(y)) | \Omega \rangle$$

Here the time ordering T means the following. By the definition,

$$T(\phi(x)\phi(y)) = \phi(y)\phi(x)$$

if $x - y \in V_-$, and

$$T(\phi(x)\phi(y)) = \phi(x)\phi(y)$$

otherwise. Since $\phi(x)$ commutes with $\phi(y)$ when $x - y$ is space-like, the function (1.1) is even. Because of the Poincare invariance axiom, (1.1) is a function of $x - y$. Denote this function by $W(x)$. It follows from the Poincare invariance that the function $W(x)$ actually depends only on x^2 .

Let \mathcal{H}_1 be the closed subspace of \mathcal{H} spanned by vectors of the form $\phi(f)\Omega$, where f is a Schwartz function on V . The space \mathcal{H}_1 is a representation of the Poincare group, and all its irreducible components have spin zero, i.e. have the form $L^2(O_m^+)$, where O_m^+ is the upper sheet of the two-sheeted hyperboloid $k^2 = m^2$ in the dual space V^* . This happens because we have a homomorphism of representations $\mathcal{S}(V) \rightarrow \mathcal{H}_1$, given by $f \rightarrow \phi(f)\Omega$, whose image is dense in \mathcal{H}_1 .

Let $\rho_1 : P \rightarrow \text{Aut}(\mathcal{H}_1)$ be the action of the Poincare group in \mathcal{H}_1 . If $x \notin V_-$, we can write $W(x)$ as

$$1.2 \quad W(x) = \lim_{\varepsilon \rightarrow 0} \langle \rho_1(x)v_\varepsilon, v_\varepsilon \rangle, \quad v_\varepsilon = \varphi(\delta_\varepsilon)\Omega,$$

where δ_ε is a family of smooth Schwartz functions tending to the δ -function.

Assume first that \mathcal{H}_1 is an irreducible representation, i.e.

$$\mathcal{H}_1 = L^2(O_m^+).$$

¹Notes by Pasha Etingof, TeXnical editing Misha Verbitsky

In this case, it is very easy to evaluate (1.2) explicitly. Indeed, the operator $\rho_1(x)$ in this case is just the operator of multiplication by the function

$$e^{i(k,x)}, \quad (k \in \mathcal{O}_m^+ \subset V^*).$$

It is easy to see that there exists a limit as $\varepsilon \rightarrow 0$ of v_ε in the sense of distributions. Because of P -invariance, this limit is an $O(V)$ -invariant distribution on \mathcal{O}_m^+ . Therefore, this limit equals to a constant function on \mathcal{O}_m^+ . We can always normalize this constant to 1, by rescaling the map φ . Then from (1.2) we will get

$$1.3 \quad W(x) = W_m(x) := \int_{\mathcal{O}_m^+} e^{i(k,x)} dk, \quad x \notin V_-; \quad W(-x) = W(x),$$

where dk is the $O(V)$ -invariant measure on \mathcal{O}_m^+ : if $k = (k_0, k_1)$, $k_0 \in \mathbb{R}$, $k_1 \in \mathbb{R}^{n-1}$, then $dk = dk_1 / \sqrt{k_1^2 + m^2}$, and dk' is the Lebesgue measure on \mathbb{R}^{n-1} .

Let δ_m^+ denote the delta function of the upper sheet of the hyperboloid $k^2 = m^2$, and δ_m^- the delta function of its lower sheet. Then (1.3) says that $W_m(x)$ equals to the Fourier transform of δ_m^+ in V_+ , of δ_m^- in V_- , and to both of them in the rest of V (these Fourier transforms are equal outside of $V_+ \cup V_-$).

It is clear from this description that it is more convenient to work with the Fourier transform of the function $W_m(x)$ than with this function itself. The Fourier transform of $W_m(x)$ is obtained by a direct computation, similar to one from Kazhdan's lectures. The answer is called the Klein-Gordon propagator:

$$1.4 \quad \tilde{W}_m(k) = w_m(k^2) = \frac{i}{k^2 - m^2 + i\varepsilon},$$

where by definition $\frac{1}{k^2 - m^2 + i\varepsilon}$ is the distribution on V^* obtained as the weak limit of $\frac{1}{k^2 - m^2 + ia}$ as $a \rightarrow +0$.

Now we turn to the general case, when \mathcal{H}_1 is not an irreducible representation of P . We will define the spectral measure μ of \mathcal{H}_1 (on \mathbb{R}) by the following rule. Let $\alpha \in \mathcal{S}(\mathbb{R})$ be a positive function. We define the integral $\int \alpha(s) d\mu(s)$ by

$$1.5 \quad \int \alpha(s) d\mu(s) = \langle \varphi(f), \varphi(f) \rangle,$$

where f is any function in $\mathcal{S}(V)$ such that $\int_{\mathcal{O}_s^+} |\tilde{f}|^2 dk = \alpha(s)$, (here \tilde{f} is the Fourier transform of f). It is easy to show that the r.h.s. of (1.4) does not depend on the choice of f , so the measure $\mu(s)$ is well defined. By the definition of $\mu(s)$, the function $W(x)$ can be expressed in the form

$$W(x) = \int W_s(x) d\mu(s).$$

Therefore, the Fourier transform of $W(x)$ has the form

$$\tilde{W}(k) = w(k^2) = \int \frac{id\mu(s)}{k^2 - s^2 + i\varepsilon}.$$

It follows from scattering theory (See Kazhdan's lectures) that if \mathcal{H} has a subrepresentation of P isomorphic to $L^2(\mathcal{O}_m^+)$ then it has components of the form $L^2(\mathcal{O}_s^+)$, occurring as continuous spectrum, for all $s \geq 2m$. Therefore, one expects that for a sufficiently generic quantum field theory, the space \mathcal{H}_1 will already contain some of this spectrum, and therefore it will be possible to see it looking at the measure μ , i.e. at the function $W(x)$.

In the case when the field theory φ is a small perturbation of the theory φ_0 of a free scalar field of mass m_0 , it is expected that \mathcal{H} has a P -invariant subspace isomorphic to $L^2(\mathcal{O}_m^+)$ for m close to m_0 , a continuous spectrum from $2m$ to ∞ , and, possibly, a finite number of discrete components. This assumption can be tested by looking at μ : it would mean that μ is supported at $\{m\} \cup [2m, \infty)$, has an atom at $s = m$, and is absolutely continuous with respect to the Lebesgue measure for $s \geq 2m$ (the new discrete spectrum makes an exponentially small contribution to μ , with respect to the deformation parameter, and so it is not seen in the perturbation expansion).

In terms of the function $w(s)$, this means the following. Let us analytically continue $w(s)$ to a complex analytic function. Then, the conjectural behavior of the spectrum that we described above would mean that $\tilde{w}(s)$ has a pole at $s = m^2$ and a cut from $4m^2$ to $+\infty$, with jump $-2\pi d\mu/ds$ when crossing the cut from up to down at the point s .

Now we will take a concrete quantum field theory and compute a perturbative expansion of the function $w(k^2)$, in order to find out if it really has such analytic properties.

1.2. The ϕ^3 -theory.

Now we consider the quantum field theory with the Lagrangian

$$1.6 \quad \mathcal{L} = \int \left(\frac{1}{2} (\nabla\phi)^2 - \frac{m^2}{2} \phi^2 + \frac{g}{3!} \phi^3 \right) d^n x.$$

This theory is a perturbation of the theory of a free scalar field of mass m with respect to a small parameter g . It is called the ϕ^3 -theory.

Remark. From physical point of view, the ϕ^3 -theory is unsatisfactory, since the energy in this theory is not bounded below, for any finite nonzero value of g . However, one can consider this theory perturbatively, i.e. regard g as an infinitesimal formal parameter, which in algebraic terms means that we work over the ring $\mathbb{C}[[g]]$. Using this theory as an example, we will do some Feynman diagrams computations, which are done in a similar manner in more complicated but more physically meaningful theories.

According to the rules of quantum mechanics, if the ϕ^3 -theory actually existed, the correlation functions of this theory would be given by the formal expression

$$1.7 \quad \langle \Omega | \phi(x_1) \dots \phi(x_N) | \Omega \rangle = \frac{1}{Z} \int \phi(x_1) \dots \phi(x_N) e^{i\mathcal{L}(\phi)} D\phi,$$

where

$$1.8 \quad Z = \int e^{i\mathcal{L}(\phi)} D\phi$$

is the partition function. As usual, these “formulas” do not apriori make sense, as the formal expression $e^{i\mathcal{L}(\phi)} D\phi$ does not represent a measure on the space of fields. However, if $g = 0$ (the free theory), we can use (1.7) as a definition: define the integral

$$1.9 \quad \int \phi(x_1) \dots \phi(x_n) e^{i\mathcal{L}(\phi)} D\phi$$

to be equal to $\langle \Omega | \phi(x_1) \dots \phi(x_N) | \Omega \rangle$. Further, for $g \neq 0$, we can expand (1.7) in a formal series in powers of g , and successive coefficients will be expressed as finite-dimensional integrals of expressions of the form (1.9). If we can compute these finite-dimensional integrals, we can get natural definition of (1.7). This computation is done using the Feynman diagrammatic techniques. Unfortunately, it turns out that some of these integrals are divergent and need to be renormalized.

Remark. Strictly speaking, in ϕ^3 -theory as we stated it, the 1-point function does not vanish. However, as we explained before, this problem can be removed by shifting the quantization map φ . In the language of functional integral, it corresponds to adding an auxiliary linear term to the Lagrangian, and in the language of Feynman diagrams it corresponds to ignoring graphs with one external vertex, and all graphs which contain such graphs as subgraphs. In fact, this shifting procedure is not so trivial, since the integral representing the 1-point function is divergent for $n \geq 1$. Thus the shifting procedure requires renormalization of some graphs with 1 external vertex. Therefore, we will come back to this topic at the end of the lecture. Until then, we will assume that the 1-point function has been normalized to zero.

1.3. Perturbative expansion of Feynman integrals

In this part of the lecture we will remind how to compute the perturbative expansion of Feynman integrals. For simplicity consider the finite-dimensional case (Kazhdan's lectures). Suppose we have a finite dimensional real vector space \mathcal{S} with a positive definite symmetric bilinear form B . Let dv be a Lebesgue measure on \mathcal{S} such that

$$1.10 \quad \int_{\mathcal{S}} e^{-B(v,v)/2} dv = 1.$$

We want to learn to compute the integral

$$1.11 \quad \int_{\mathcal{S}} P(v) e^{-B(v,v)/2} dv,$$

where $P : \mathcal{S} \rightarrow \mathbb{R}$ is a polynomial. This integral is a sum of integrals of the form

$$1.12 \quad \langle f_1 \dots f_N \rangle_0 := \int_{\mathcal{S}} f_1(v) \dots f_N(v) e^{-B(v,v)/2} dv,$$

where $f_1, \dots, f_N \in \mathcal{S}^*$. It is clear that if N is odd, this integral vanishes, as the integrand is an odd function. Thus, it is enough to consider the case when $N = 2K$. In this case, the answer is given by the following formula, which is called the Wick formula.

Proposition 1.1

$$1.13 \quad \langle f_1 \dots f_{2N} \rangle_0 = \sum_{s \in S_{2K}/\sim} B^{-1}(f_{s(1)}, f_{s(2)}) \dots B^{-1}(f_{s(2K-1)}, f_{s(2K)}),$$

where B^{-1} is the inverse form to B on \mathcal{S}^* , S_{2K} is the symmetric group, and $s_1 \sim s_2$, $s_1, s_2 \in S_{2K}$ if they define the same term in (1.13).

The proof is obvious: the right hand side of (1.13) is the only polylinear expression in f_i invariant under the group $O(\mathcal{S}) \times S_{2K}$, up to a factor, and the normalization is deduced from the case $f_1 = \dots = f_{2K} = f$.

Remark. The set S_{2K}/\sim is the set of pairings, i.e. splittings of the set $\{1, \dots, 2K\}$ into pairs. Thus, terms of the r.h.s. of (1.13) are in 1-1 correspondence with pairings. In particular, the number of these terms is equal to $(2m)!/2^m m!$.

Now consider a perturbation of this situation. Let

$$Q(v) = \sum_{n \geq 1} g_n Q_n(v^{\otimes n})/n!,$$

where g_m are formal variables and $Q_m \in \mathcal{S}^m \mathcal{S}^*$. Consider the integral

$$1.14 \quad \langle f_1 \dots f_N \rangle = \int_{\mathcal{S}} f_1(v) \dots f_N(v) e^{-B(v,v)/2 + Q(v)} dv \in \mathbb{C}[[g_1, g_2, g_3, \dots]].$$

This integral can be computed using Feynman graphs, as follows.

Let $\mathfrak{n} = \{n_1, n_2, n_3, n_4, \dots\}$ be any sequence of nonnegative integers which is eventually zero. Let $G(N, \mathfrak{n})$ be the set of equivalence classes of all graphs which have N 1-valent vertices labeled by $1, \dots, N$, and n_i unlabeled i -valent vertices, $i \geq 1$. The labeled vertices are called external, and the unlabeled ones internal.

Any graph $\Gamma \in G(N, \mathfrak{n})$ defines a polylinear function F_Γ of f_1, \dots, f_N which is defined as follows. At the 1-valent vertex of Γ labeled by i one places the vector f_i , at every m -valent vertex one places the tensor Q_m , and then takes contractions along edges using the form B^{-1} (see Kazhdan's lectures). Then one has

Theorem 1.2

$$1.15 \quad \langle f_1, \dots, f_N \rangle = \sum_{\mathfrak{n}} \prod_i g_i^{n_i} \sum_{\Gamma \in G(N, \mathfrak{n})} |\text{Aut}(\Gamma)|^{-1} F_\Gamma(f_1, \dots, f_N),$$

where $\text{Aut}(\Gamma)$ denotes the group of automorphisms of Γ which fix the external vertices.

The proof is easy. First one observes that (1.13) is a special case of (1.15), when $\mathfrak{n} = 0$. Next, one can think of each i -valent vertex of a graph Γ as a collection of i 1-valent vertices which are situated very close to each other. Then it is clear that any graph $\Gamma \in G(N, \mathfrak{n})$ defines as many as $|\text{Aut}(\Gamma)|^{-1} \prod_i i^{n_i} \prod n_i!$ different pairings of such 1-valent vertices. Thus, formula (1.13) implies (1.15).

1.4. Computation of a Feynman integral over functions on a Minkowski space.

Now we will try to compute the function $w(k^2)$. Using formula (1.7), we get

$$1.16 \quad W(x) = Z^{-1} \int T(\phi(x)\phi(0)) e^{i\mathcal{L}(\phi)} D\phi$$

Applying formula (1.15), we express the r.h.s. of (1.16) as a sum over graphs. The function $Q(\phi)$, which was considered above for the finite dimensional case, is of the form $\frac{g}{3!} Q_3(\phi)$, where Q_3 is a cubic form given by

$$Q_3(\phi) = \langle \phi(a) \otimes \phi(b) \otimes \phi(c), i\delta_{a=b=c} \rangle.$$

Thus, we get a sum over all graphs with two external 1-valent vertices, and a number of internal trivalent vertices.

Graphs which have components without external vertices will not occur in this sum, since we have divided by Z . So there are two remaining types of graphs: connected and disconnected. Disconnected graphs have two components, each having one external vertex.

It is easy to see that the sum over all disconnected graphs equals

$$\langle \Omega | \phi(0) | \Omega \rangle^2 = 0.$$

Therefore, disconnected graphs can be ignored.

We see that the function $w(k^2)$, can be written as a sum over connected graphs with two external vertices. So, we can represent $w(k^2)$ in the form

$$1.17 \quad w(k^2) = \sum_{j=0}^{\infty} w^{(j)}(k^2),$$

where $w^{(j)}(k^2)$ is the sum over all connected graphs which are chains of j 1-particle irreducible graphs. We have $w^{(0)}(k^2) = w_m(k^2) = \frac{i}{k^2 - m^2 + i\epsilon}$ (the free propagator), and

$$1.18 \quad w^{(j)}(k^2) = w^{(0)}(k^2) \left(\frac{w^{(1)}(k^2)}{w^{(0)}(k^2)} \right)^j.$$

Set

$$1.19 \quad \Sigma(k^2) = i \frac{w^{(1)}(k^2)}{w^{(0)}(k^2)^2}.$$

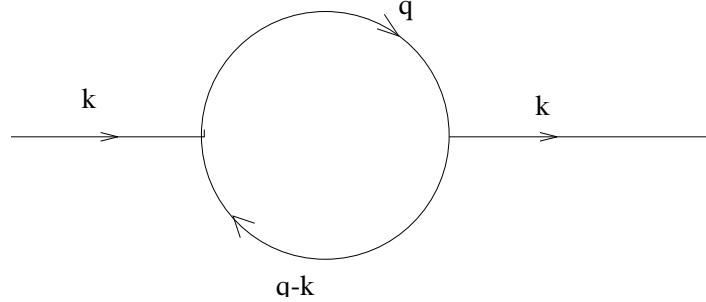
(It is easy to see that $\Sigma(k^2)$ is a real function). Summing the geometric progression in (1.17), we get

$$1.19 \quad w(k^2) = \frac{i}{k^2 - m^2 - \Sigma(k^2) + i\varepsilon}.$$

Thus, it remains to compute $\Sigma(k^2)$. We hope to find that $\Sigma(k^2)$ is analytic near $k^2 = m^2$, and has a cut at $k^2 \geq 4m^2$. This would confirm the expected analytic behavior of $w(k^2)$.

In our setup, the bilinear form B^{-1} from the previous section has the kernel $\frac{i}{(k_1 - k_2)^2 - m^2 + i\varepsilon}$, and the 3-tensor Q_3 has the kernel $i\delta(k_1 + k_2 + k_3)$. Therefore, the term corresponding to each graph Γ in the sum for $\Sigma(k^2)$ is an integral of a rational function over the space $H^1(\Gamma, \mathbb{R}) \otimes V^*$. So the number of integrations is proportional to the number of loops in the graph. Therefore, these integrals, over loop parameters, are called loop integrals.

Now we will compute the first nontrivial coefficient of the g -expansion of Σ . We have $\Sigma = g^2 \Sigma_2 + O(g^3)$, where Σ_2 is the sum over all connected graphs with two trivalent vertices (there is no graph with only one trivalent vertex). It is easy to see that there is only one such graph, namely the following one:



So we must put an integration parameter $q \in V^*$ on the only loop of the graph. Thus, the function $\Sigma_2(k^2)$ can be represented as a single loop integral

$$1.19 \quad \Sigma_2(k^2) = \frac{i}{2} \int_V \frac{d^n q / (2\pi)^n}{(q^2 - m^2 + i\varepsilon)((q - k)^2 - m^2 + i\varepsilon)}.$$

(division by 2 arises from the automorphism group of Γ , which has order 2).

Remark 1. When we compute the group of automorphisms of a graph, we do not take into account the orientation of edges. The arrows are put on edges arbitrarily, in order to balance the momenta which are attached to edges. The distribution of momenta should satisfy the condition that the sum of incoming momenta at each internal vertex should equal the sum of outgoing ones.

Remark 2. It may appear that the integral in (1.19) is real, but in fact it is not. It is easy to see that for real negative k^2 the integral is imaginary, so the r.h.s. of (1.19) is real. This happens because of addition of $i\varepsilon$ in the denominator.

Now we try to compute integral (1.19). The integrand is a fraction where the denominator is a product of two different factors. For a general graph, the integrand will have a product of many different factors in the denominator, which causes inconvenience. There is a remarkable trick, which

allows to convert this integral into an integral of a function whose denominator is a power of a single factor. This trick is the ‘‘Feynman famous formula’’:

Proposition 1.3

$$1.20 \quad \frac{1}{A_1 \dots A_N} = \int_{\alpha_1 + \dots + \alpha_N = 1} \frac{1}{(\sum \alpha_i A_i)^N} d\sigma,$$

where $d\sigma$ is a Lebesgue measure on the simplex with volume 1.

Proof We prove the statement by induction in N . For $N = 1$, the statement is obvious. Let $N > 1$. Denote the r.h.s. of (1.20) by $I_N(A_1, \dots, A_N)$. Then from a homogeneity argument it follows that

$$1.21 \quad \int_{s \leq \alpha_1 + \dots + \alpha_N \leq t} \frac{1}{(\sum \alpha_i A_i)^N} d\alpha_1 \wedge \dots \wedge d\alpha_N = \frac{1}{(N-1)!} I_N(A_1, \dots, A_N) \ln(t/s).$$

Now observe that the N -form under the integral in (1.21) has the form $d\omega$, where

$$1.22 \quad \omega = \sum_{j=1}^N \frac{(-1)^j}{N(N-1)A_j(\sum \alpha_i A_i)^{N-1}} d\alpha_1 \wedge \dots \wedge \widehat{d\alpha_j} \wedge \dots \wedge d\alpha_N.$$

Therefore, using Stokes’ formula, we get

$$1.23 \quad \begin{aligned} & \sum_{j=1}^N \frac{1}{N(N-1)A_j} \frac{1}{(N-2)!} I_{N-1}(A_1, \dots, \hat{A}_j, \dots, A_N) \ln(t/s) \\ &= \frac{1}{(N-1)!} I_N(A_1, \dots, A_N) \ln(t/s). \end{aligned}$$

(the integrals of ω over simplices $\sum \alpha_i = s, t$ cancel each other, and only the integrals over N truncated simplices remain). Using the induction assumption, we obtain (1.20). \square

We will use formula (1.20) for $N = 2$. In this case it has the form

$$1.24 \quad \frac{1}{AB} = \int_0^1 \frac{d\alpha}{(\alpha A + (1-\alpha)B)^2}.$$

Applying this formula to (1.19), and shifting the integration variable, we obtain

$$1.25 \quad \Sigma_2 = \frac{i}{2} \int_0^1 \int_V \frac{1}{(q^2 + \alpha(1-\alpha)k^2 - m^2 + i\varepsilon)^2} \frac{d^n q}{(2\pi)^n} d\alpha.$$

It is convenient now to perform a Wick rotation of the integration cycle. Write any vector $q \in V$ as (q_0, q_1) , $q_0 \in \mathbb{R}$, $q_1 \in \mathbb{R}^{n-1}$, so that $q^2 = q_0^2 - q_1^2$ (where q_1^2 is the usual squared norm). Consider the 1-parameter family of integration cycles, $C(t) = \{(e^{\pi i t/2} q_0, q_1) \in V_{\mathbb{C}} | (q_0, q_1) \in V\}$, $0 \leq t \leq 1$. The integral (1.25) is over $C(0)$. It is clear that during the deformation $C(t)$ we do not pick up any poles of the integrand in (1.25) (since $\varepsilon > 0$). Therefore, if (1.25) converges, the integral over $C(0)$ (which is Σ_2) equals to the integral over $C(1)$.

The cycle $C(1)$ is a real subspace of $V_{\mathbb{C}}$ which carries a natural positive definite metric, $|q|^2 = -q^2$. Thus, integral (1.25) can be written as

$$1.26 \quad \Sigma_2(k^2) = \frac{i}{2} \int_0^1 \int_{\mathbb{R}^d} \frac{1}{(|q|^2 - \alpha(1-\alpha)k^2 + m^2)^2} \frac{d^n q}{(2\pi)^n} d\alpha.$$

($i\varepsilon$ is no longer necessary, as the integrand is now smooth for negative k^2).

It is now an exercise to compute this integral explicitly in elementary functions. However, we are more interested in qualitative properties of the answer. Namely, from formula (1.26) it is obvious that

1) Σ_2 is analytic near $k^2 = m^2$.

2) Σ_2 has a cut at the set of those values of k^2 for which the integrand is singular.

Since the function $\alpha(1 - \alpha)$ varies between 0 and 1/4 as α varies from 0 to 1, the cut occurs at $k^2 \geq 4m^2$. Thus, the function Σ_2 has the expected analytic behavior.

Remark. We could have considered from the very beginning a Euclidean theory. This means, V is a Euclidean space, the Lagrangian is

$$1.27 \quad \int \left(\frac{1}{2} (\nabla\phi)^2 + \frac{1}{2} m^2 \phi^2 + \frac{g}{3!} \phi^3 \right) d^n x,$$

The correlation functions are

$$1.28 \quad \langle \phi(x)\phi(0) \rangle = \int_{S(V)} \phi(x)\phi(0) e^{-\mathcal{L}(\phi)} D\phi,$$

the propagator is $\frac{1}{k^2+m^2}$, the cubic functional corresponding to a trivalent vertex is $-\frac{1}{3!}\delta(k_1+k_2+k_3)$, and the Fourier transform of (1.28) is $\frac{1}{k^2+m^2+\Sigma}$, where $-\Sigma$ is computed as a sum over 1-particle irreducible connected graphs. If we try to compute Σ_2 , we will get exactly the same answer as given by (1.26), i.e. $\Sigma_2^{\text{Euclidean}}(|k|^2) = \Sigma_2^{\text{Minkowski}}(-k^2)$. This is a general phenomenon: the Wick rotated answer for a theory in the Minkowski space coincides with the answer for the corresponding Euclidean theory.

1.5. Renormalization of divergent graphs.

Unfortunately, the integrals (1.25), (1.26) are divergent for $n \geq 4$, since the integrand does not decay rapidly enough at infinity. Renormalization techniques help to give meaning to these integrals anyway.

From now until the end of the lecture we consider the case $n = 4$, and the Euclidean picture. In this case, the easiest way to make sense of (1.26) is to differentiate with respect to k^2 . If we differentiate under the integral sign, we will obtain a convergent integral, and then we can integrate it back, which will define Σ_2 up to a constant. However, this is, in general, not the best way to proceed. The method which is usually applied in renormalization theory is the following.

We will replace the propagator $\frac{1}{k^2+m^2}$ with a more rapidly decreasing propagator depending on a parameter Λ , of the form $P_\Lambda = \frac{\chi(k^2, \Lambda^2)}{k^2+m^2}$, where $\chi(k^2, \Lambda^2)$ is a smooth function with sufficiently rapid decay at $k^2 \rightarrow \infty$ for a fixed Λ , which tends to 1 at $\Lambda \rightarrow \infty$ (it is called the cutoff function). For instance, one can take $\chi(k^2, \Lambda^2) = \frac{\Lambda^4}{(k^2+\Lambda^2)^l}$, $l \in \mathbb{N}$. Computing Σ_2 for this new propagator, we will obtain a convergent integral for each finite value of Λ , and the answer will depend on Λ in the following way:

$$1.29 \quad \Sigma_2(k^2, \Lambda, m) = -A \ln(\Lambda/m) + O(1), \Lambda \rightarrow \infty,$$

where A is a positive constant. Of course, the limit of (1.29) as $\Lambda \rightarrow \infty$ (which would be the value of (1.26)) does not exist: one says that the integral is logarithmically divergent.

Now fix a constant $m = m_R > 0$ (the renormalized mass). Given Λ , we will adjust the mass m_0 of the theory in such a way that the Σ -function $\Sigma(k^2, \Lambda, m_0, g)$ for the theory with this mass and the propagator P_Λ has a pole exactly at $k^2 = -m_R^2$. That is, define $m_0(m_R, g, \Lambda) = m_R + O(g)$ by the equation

$$m_R^2 = m_0^2 + \Sigma(-m_R^2, \Lambda, m_0, g).$$

It is clear the solution of this equation (modulo g^3) has the form $m_0^2 = m_R^2 - g^2 \Sigma_2(-m_R^2, \lambda, m_0)$, so by (1.29) it has the form $m_R^2 + g^2 A \ln(C\Lambda/m_R)$, where C is a constant which depends only on the cutoff function χ . It is easy to see that there exists a limit

$$\Sigma_2^R(k^2, m_R, g) = \lim_{\Lambda \rightarrow \infty} \Sigma(k^2, \Lambda, m_0(m_R, g, \Lambda), g).$$

This limit is called the renormalized Σ -function. Of course, it depends on the choice of m_R .

Let us now understand what this renormalization does to the Feynman diagram expansion. The new Lagrangian of the theory, with the mass m_0 , differs from the old one (with mass $m = m_R$) by the additional quadratic term $\frac{1}{2} A g^2 \ln(C\Lambda/m_R) \phi^2$. Therefore, according to the rules of Feynman calculus, the g^2 -term of the perturbation expansion will now be the sum of terms corresponding to two Feynman diagrams:



At the 2-valent vertex of the second diagram, we put the tensor

$$-A g^2 \ln(C\Lambda/m_R) \delta(k_1 - k_2).$$

This gives us an extra constant summand in $\Sigma_2(k^2, \Lambda, m_0)$, which compensates the divergence and ensures that there exists a limit Σ_2^R of

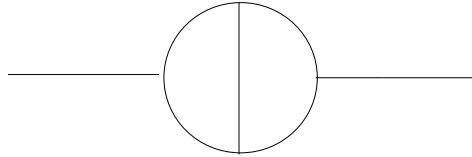
$$\Sigma_2(k^2, \Lambda, m_0(m_R, g, \Lambda))$$

as $\Lambda \rightarrow \infty$.

Remark. We have chosen our renormalization in such a way that the function $w(k^2)$ (modulo g^3) has a pole at $k^2 = -m_R^2$. This is the reason that the constant m_R is called the renormalized mass.

1.6. Renormalization in higher orders.

Now let us consider the terms our Feynman diagrams expansion which come with a power of g higher than 2. For example, consider the following graph,



which occurs with g^4 . It is easy to see that this graph defines a convergent integral for $n = 4$. We will see that this is also the case for more complicated graphs. Namely, for $n = 4$, the more vertices a graph has, the better is the rate of convergence of the corresponding integral.

Let us analyse the situation for arbitrary n . It is convenient to introduce the following definitions.

Definition 1 *The superficial divergence index $div(\Gamma)$ of a graph Γ is the difference of the degrees of the numerator and denominator in the integrand of the corresponding integral.*

Remark. By the definition, the degree of the integration variables q_i and their differentials dq_i equals 1.

Definition 2 *A graph is called superficially divergent if $div(\Gamma) \geq 0$ and superficially convergent if $div(\Gamma) < 0$.*

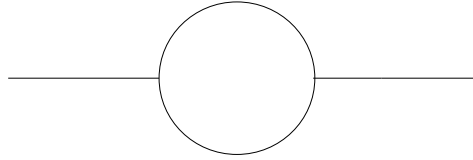
Remark. Γ is called logarithmically, linearly, quadratically,... divergent if $div(\Gamma) = 0, 1, 2, \dots$

Proposition 1.4 In the ϕ^3 -theory, $div(\Gamma) = (n - 6)b_1 + 6 - 2E$, where b_1 is the number of loops, and E the number of external vertices.

Proof Let M be the number of internal edges, and N the number of external vertices. Since Γ is connected and has only trivalent internal vertices, we have $M = \frac{3N-E}{2}$, $b_1 = \frac{N+2-E}{2}$. It is easy to see that the degree of the numerator in the integral corresponding to Γ is nb_1 (b_1 loop integrations over \mathbb{R}^n), and the degree of the denominator is $2M$ (M quadratic factors). Thus, $div(\Gamma) = nb_1 - 2M = (n-6)b_1 + 6 - 2E$.

Proposition 1.4 implies the following.

- 1) If $n > 6$, divergence worsens as the number of vertices grow.
- 2) If $n = 6$, all graphs with $E = 2, 3$ are equally bad (have $div(\Gamma) = 6 - 2E$), while for $E \geq 4$ they are superficially convergent.
- 3) For $n \leq 3$, all graphs with $E \geq 2$ are superficially convergent.
- 4) For $n = 4$, the only superficially divergent graph with $E \geq 2$ is Γ_2 :

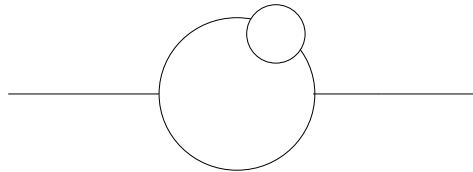


- 5) For $n = 5$, there are finitely many superficially divergent graphs.

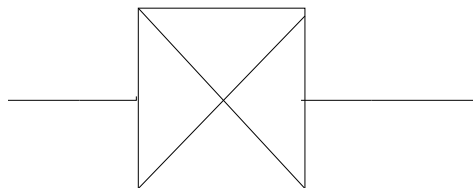
It is clear that superficial convergence of a graph is necessary but not sufficient for convergence of the corresponding integral. Indeed, a superficially convergent integral may have a subintegral that diverges. However, one can formulate a sufficient condition for actual convergence in terms of superficial convergence. This condition is given by Weinberg theorem.

Weinberg theorem Let Γ be a graph such that the integral of the corresponding function over any subset of the set of loops of Γ is superficially convergent. Then the integral corresponding to Γ is convergent.

Weinberg theorem guarantees that in the ϕ^3 -theory all integrals are convergent for $n \leq 3$. Let us see what happens if $n = 4$. In this case, we have one superficially divergent graph Γ_2 . Of course, there are infinitely many superficially convergent but still divergent graphs, namely, all graphs which contain Γ_2 as a subgraph, e.g.



However, we have renormalized the graph Γ_2 , i.e. compensated its divergence by adding another auxiliary graph of the form



As a result, all graphs which are divergent because they contain Γ_2 will be renormalized automatically and become convergent. That is, the divergence of each graph containing Γ_2 will be compensated by a counterterm, in which Γ_2 will be replaced by the above auxiliary graph. The fact that this ensures convergence in all orders follows from the “Strong Weinberg theorem”, which states, roughly, that if all graphs at orders $\leq N$ (in g) have been renormalized, then all superficially convergent graphs at order $N+1$ are actually convergent. Thus, after mass renormalization all correlation functions become well defined.

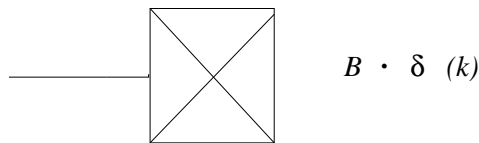
Remark. As we explain above, the procedure of setting the 1-point function to zero in ϕ^3 -theory also requires renormalization. Namely, we have two graphs with one external vertex,



among which the first is quadratically divergent and the second logarithmically. Consider a new Lagrangian of the form $\mathcal{L}' = \mathcal{L} + B(m, g, \Lambda) \int \phi$, where $B(m, g, \Lambda) = a(m, g)\Lambda^2 + b(m, g)\Lambda + c(m, g) \ln(\Lambda/m) + d(m, g)$, and choose the constants a, b, c, d in such a way that the limit of the 1-point function for this Lagrangian, computed for the cutoff propagator at Λ , tends to 0. The constants a, b, c, d are determined by this condition uniquely, but they of course depend on the cutoff function χ . In the language of Feynman graphs, this corresponds to compensating the divergence in the sum



by adding a third summand



where at the vertex we put the linear functional $B \cdot \delta(k)$. This reduces the problem of renormalization of the ϕ^3 theory to renormalization of the graph Γ_2 , which was done above.

Lecture 2: Perturbative renormalization (continued)

Edward Witten²

October 1996

2.1. Renormalizability of quantum field theories.

Last week we considered the ϕ^3 theory and found that its renormalizability (i.e. how many graphs are divergent, and how badly divergent they are) depends on whether $n < 6$, $n = 6$, or $n > 6$. Now we will see how to find that the critical value of n in the ϕ^3 theory is 6 without considering any graphs at all.

In order to find the critical value of n , we will define and compute the homogeneity degree of all terms in the Lagrangian. We recall that the Lagrangian is

$$(2.1) \quad \int \left(\frac{1}{2} (\nabla\phi)^2 + \frac{1}{2} m^2 \phi^2 + \frac{g}{3!} \phi^3 \right) d^n x,$$

(we are considering the Euclidean theory; in the Minkowski version, the story is the same). Consider the dilation $x \rightarrow t^{-1}x$, $t \in \mathbb{R}$, $x \in V$. Under this transformation $d^n x \rightarrow t^{-n} d^n x$, $\nabla \rightarrow t\nabla$. If we want the integral (2.1) to be invariant under this dilation, we must impose the following three transformation laws: $(\nabla\phi)^2 \rightarrow t^n (\nabla\phi)^2$, $m^2 \phi^2 \rightarrow t^n m^2 \phi^2$, $g\phi^3 \rightarrow t^n g\phi^3$. From the first law we get $\phi \rightarrow t^{\frac{n-2}{2}} \phi$, from the second $m \rightarrow tm$, and from the third $g \rightarrow t^{\frac{6-n}{2}} g$. In the future we will write these scaling laws like this:

$$(2.2) \quad [x] = -1, [d^n x] = -n, [\nabla] = 1, [\phi] = \frac{n-2}{2}, [m] = 1, [g] = \frac{6-n}{2}.$$

In other words, any homogeneous quantity a scales as $a \rightarrow t^{[a]}a$ under the dilation, and the number $[a]$ is called the dimension of a .

Remark. It is not difficult to define the notion of dimension completely formally (using the tensor calculus on the tangent bundle of V), but it is more illuminating to illustrate it on examples.

We see that the number 6 appears in the scaling law for g . More precisely, we see that the behavior of the ϕ^3 theory is determined by whether $[g]$ is positive, zero, or negative. Namely, we have three cases.

1. $[g] > 0$ ($n < 6$): there is a finite number of superficially divergent graphs (by graphs we always mean connected graphs).
2. $[g] = 0$ ($n = 6$): the number of superficially divergent graphs is infinite, but $div(\Gamma)$ is bounded from above.
3. $[g] < 0$ ($n > 6$): there are infinitely many graphs with any number of external vertices and arbitrarily high $div(\Gamma)$.

Now we will explain why this method works, and prove that it can be used to compute the critical value of the dimension of the spacetime in a general field theory. In general, we can consider the following setup. Suppose we have a quantum field theory with fields ϕ_1, \dots, ϕ_N , and the Lagrangian

$$(2.3) \quad \mathcal{L} = \int \left(\sum_i Q_i(\phi_i) + \sum_k g_k I_k(\phi_1, \dots, \phi_N) \right) d^n x,$$

where Q_i is a free (quadratic) part, and I_k are interaction (coupling) terms (differential monomials in fields, cubic and higher). Each interaction term comes with a small parameter g_k , which is called the coupling constant.

²Notes by Pasha Etingof and David Kazhdan, TeXnical editing Misha Verbitsky

Definition An interaction I_k is called *subcritical* if $[g_k] > 0$, and *critical* if $[g_k] = 0$. A field theory is called *superrenormalizable* if all interaction terms in the Lagrangian are subcritical, and is called *renormalizable*, or *critical*, if all interaction terms are critical or subcritical, but not all of them are subcritical. Otherwise, the theory is called *non-renormalizable*.

For example, the ϕ^3 theory is superrenormalizable for $n < 6$, renormalizable for $n = 6$, and non-renormalizable for $n > 6$.

We will assume that our theory is a perturbation of a free theory, (for the definition of a free theory, see Kazhdan's lectures). In a free theory, one can easily see that the dimension of bosonic fields is $\frac{n-2}{2}$, and of fermionic fields is $\frac{n-1}{2}$. Since dimensions of fields are determined from the quadratic part of the Lagrangian, these dimensions will be the same in the perturbed (classical) theory as well.

We will also assume that $n \geq 2$ (In the quantum mechanical case $n = 1$ renormalization theory is not necessary). Then $[\phi_i] \geq 0$.

Theorem 2.1 (i) If a theory is superrenormalizable, there is a finite number of superficially divergent graphs in its Feynman diagram expansion.

(ii) If a theory is renormalizable then the number of superficially divergent graphs is infinite, but $\text{div}(\Gamma)$ is bounded from above.

(iii). If a theory is non-renormalizable then there are infinitely many graphs with any number of external vertices and arbitrarily high $\text{div}(\Gamma)$.

Proof Let Γ be a graph in the Feynman diagram expansion of our theory. Types of internal vertices of such a graph correspond to interaction terms in the Lagrangian, and types of its external vertices and edges correspond to fields ϕ_i . Let e_i be the number of external vertices of type ϕ_i , and v_k be the number of internal vertices of type I_k . Then it is easy to show that

$$(2.4) \quad \text{div}(\Gamma) = n - \sum_i e_i [\phi_i] - \sum_k v_k [g_k].$$

Statements (i)-(iii) of the theorem follows immediately from (2.4).

Definition A theory of the form (2.3) is called *classically scale invariant* if the n -form under the integral in the Lagrangian is invariant under dilations.

For example, the theory of a free scalar field is scale invariant iff it is massless.

It is clear that a massless theory is scale invariant if and only if it is purely critical, i.e. $[g_k] = 0$ for all k .

Remark. Scale invariant theories are always conformal. Indeed, the Lagrangian is always written naturally in terms of the metric on the spacetime, and depends only on the fields and their first derivatives. This implies that a scale invariant Lagrangian has to be invariant under conformal changes of metric.

2.2. Critical dimensions of some field theories.

Now we will compute the critical dimension for several important field theories.

Example 1. Sigma-models.

Let M be a Riemannian manifold. The Lagrangian of the sigma-model on \mathbb{R}^n with the target space M is

$$(2.5) \quad \mathcal{L}(\phi) = \int d^n x g_{ij}(\phi) \nabla \phi^i \cdot \nabla \phi^j.$$

Since this theory is conformal in two dimensions, 2 must be the critical dimension. Let us show by a direct computation that this theory is not renormalizable for $n > 2$.

Let ϕ_i be coordinates on M near some point. If the metric g_{ij} is constant in these coordinates, the theory is free. Consider a nonconstant metric of the form

$$(2.6) \quad g_{ij}(\phi) = \delta_{ij} + a_{ijk}\phi^k + r_{ijkl}\phi^k\phi^l + \dots$$

(If we chose normal coordinates, we could get rid of a , but not of r , as r in normal coordinates is the Riemann curvature tensor). Substituting this into the Lagrangian, we find that $[r] = 2 - n$. This shows that sigma-model is not renormalizable beyond $n \geq 2$, unless the metric is flat.

Example 2. Gravity.

In the theory of gravity the spacetime is the space \mathbb{R}^n with a metric of the form $g_{ij} = \delta_{ij} + h_{ij} \sqrt{G}$, where G is the Newton's constant (for us, it is just a formal parameter). The Lagrangian of the theory is

$$(2.7) \quad \mathcal{L}(g) = \frac{1}{G} \int R(g) d^n x,$$

where $R(g)$ is the scalar curvature of the metric. In terms of h , this Lagrangian can be rewritten as

$$(2.8) \quad \mathcal{L} = \int d^n x ((\nabla h)^2 + \sqrt{G} h (\nabla h)^2 + \dots),$$

so we get $[h] = \frac{n-2}{2}$, $[G] = 2 - n$. Thus, as in the previous example, the theory is non-renormalizable for $n > 2$ and critical for $n = 2$.

Remark. Here $h(\nabla h)^2$ stands for an expression which is linear in h and quadratic in ∇h . It is easy to compute what it is exactly, but it does not matter to us, since we are only interested in the dimension. So we will use such sloppy notation.

Example 3. Gauge theory.

In gauge theory fields are connections in a fixed principal G -bundle on the space \mathbb{R}^n , where G is a compact Lie group. The Lagrangian has the form

$$(2.9) \quad \mathcal{L}(\tilde{A}) = \frac{1}{e^2} \int \text{Tr}(F_{\tilde{A}} \wedge *F_{\tilde{A}}),$$

where $F_{\tilde{A}}$ is the curvature of the connection \tilde{A} , and Tr is an invariant nondegenerate bilinear form on the Lie algebra \mathfrak{g} of G .

In the computation of dimension, we will assume that our G -bundle is trivial, so a connection is represented by a 1-form $A: \tilde{A} = d + A$. Then $F = dA + A \wedge A$.

Consider the field $B = A/e$. In terms of B , the Lagrangian takes the form

$$\mathcal{L} = \int ((\nabla B)^2 + eB^2 \nabla B + e^2 B^4) d^n x$$

We have $[B] = \frac{n-2}{2}$, so $[B^2 \nabla B] = \frac{3n}{2} - 2$, and $[e] = \frac{4-n}{2}$. Thus, if the group G is noncommutative, the theory is superrenormalizable for $n < 4$, renormalizable for $n = 4$, and non-renormalizable for $n > 4$. In $n = 4$, the theory is conformal (If G is commutative, the theory is free).

Remark. Dimensions of fields in our computation agree with geometric dimensions. Indeed, in geometry, since $d \rightarrow td$ under the dilation $x \rightarrow t^{-1}x$, we have $[d] = 1$, so we must have $[A] = 1$. This coincides with our result: $[A] = [Be] = [B] + [e] = \frac{n-2}{2} + \frac{4-n}{2} = 1$.

Example 4. Gauge theory with a scalar bosonic field or with a fermionic field.

Consider the setting of Example 3, and let V be a finite-dimensional representation of G . Let ϕ be a section of the corresponding vector bundle on \mathbb{R}^n (associated with the above G -bundle). The Lagrangian of the gauge theory with ϕ is

$$(2.10) \quad \mathcal{L}'(\tilde{A}, \phi) = \mathcal{L}(\tilde{A}) + \int (\nabla_{\tilde{A}}\phi)^2 d^n x,$$

where $L(\tilde{A})$ is the Lagrangian (2.9). Writing \tilde{A} in the form $\tilde{A} = d + A$, $A = eB$, we get

$$\mathcal{L}' = \int ((\nabla B)^2 + eB^2\nabla B + e^2B^4 + (\nabla\phi)^2 + 2e(\nabla\phi, B\phi) + e^2(B\phi)^2) d^n x$$

For this formula we get $[B] = [\phi] = \frac{n-2}{2}$, and (even when G is commutative!) $[e] = \frac{4-n}{2}$. Thus, the theory is superrenormalizable in $n < 4$, renormalizable in $n = 4$, and nonrenormalizable for $n > 4$ (regardless of the commutativity of G).

The same answer applies if we have a fermionic field. Let ψ be a section of $V \otimes S$, where S is the spin bundle over the spacetime. The Lagrangian of the gauge theory with ψ is

$$(2.11) \quad \mathcal{L}''(\tilde{A}, \psi) = \mathcal{L}(\tilde{A}) + \int (\psi, D_{\tilde{A}}\psi) d^n x,$$

where $D_{\tilde{A}}$ is (i times) the Dirac operator along the connection \tilde{A} . The critical dimension in this theory, as before, is 4, for any nontrivial compact group G .

Remark. If $G = U(1)$ and V is the standard 1-dimensional representation of G , this theory is the quantum electrodynamics (QED).

Example 5. Theory of a scalar bosonic field.

Consider the Lagrangian

$$\mathcal{L}(\phi) = \int \left(\frac{1}{2}(\nabla\phi)^2 + \frac{m^2}{2}\phi^2 + Q(\phi) \right) d^n x,$$

where $Q(\phi) = \sum g_k \phi^k$. We already considered this type of Lagrangian in Lecture 1. In particular, the ϕ^3 -theory is a special case of this situation, when Q is a cubic polynomial. We have $[g_k] = n - k\frac{n-2}{2}$. Thus, for $n = 2$ all terms in the Taylor expansion of $Q(\phi)$ are subcritical. For $n = 3$, the term ϕ^k is subcritical for $k < 6$, critical for $k = 6$, and non-renormalizable for $k > 6$. For $n = 4$, the term ϕ^k is subcritical for $k < 4$, critical for $k = 4$, and non-renormalizable for $k > 4$. For $n = 5, 6$, the terms ϕ^k , $k \geq 4$ are non-renormalizable, and the term ϕ^3 is subcritical for $n = 5$ and critical for $n = 6$.

In particular, the following theories are critical: the ϕ^3 theory in 6 dimensions, the ϕ^4 -theory in 4 dimensions, and the ϕ^6 theory in 3 dimensions.

Example 6. Yukawa interaction.

Consider a theory with a scalar bosonic field ϕ and a fermionic field ψ . We will consider the Lagrangian

$$(2.12) \quad \mathcal{L}(\phi, \psi) = \int ((\nabla\phi)^2 + (\psi, D\psi) + g\phi[\psi, \psi]) d^n x$$

(ψ takes values in a vector bundle W which is a direct sum of several copies of the spin bundle; (\cdot, \cdot) , $[\cdot, \cdot]$ are a symmetric and a skew-symmetric form on W which are invariant under gauge transformations). The cubic term $\phi[\psi, \psi]$ is called the Yukawa interaction. Let us compute the dimension of its coefficient g .

We have $[\phi] = \frac{n-2}{2}$, $[\psi] = \frac{n-1}{2}$, so $[g] = \frac{4-n}{2}$. So this theory is critical in dimension 4, superrenormalizable in dimension < 4 , and non-renormalizable in dimension > 4 .

We observe that in dimension 4 all interactions except ϕ^3 , ϕ^4 and $\phi[\psi, \psi]$ are “bad” (non-renormalizable).

Example 7. Standard model.

Let us now try to write down the most general renormalizable theory that lives in our 4-dimensional physical world. According to the above examples, we cannot include gravity or Σ -model, but we can include connections, bosons with terms up to degree 4, and fermions with Yukawa interaction. If we only take these fields, these are the only renormalizable terms we can write. Thus the most general Lagrangian we can write in dimension 4 giving a renormalizable theory is

$$(2.13) \quad \mathcal{L}(A, \phi, \psi) = \int (e^{-2} F_A^2 + (\psi, D\psi) + (\nabla\phi)^2 + g_1\phi^4 + g_2\phi\psi^2 + \text{lower terms}) d^4x.$$

The Standard Model is a theory which belongs to this family, with the group G containing $SU(3) \times SU(2) \times U(1)$.

2.3. Perturbative renormalization of critical theories.

From now on we will consider only critical theories. As a model example we can consider ϕ^4 -theory in 4 dimensions, which has one critical interaction ϕ^4 , or its extension containing fermions, which has an additional critical interaction $\phi\psi^2$ (the Yukawa interaction).

Consider the Lagrangian of ϕ^4 -theory:

$$(2.14) \quad \mathcal{L} = \int \left(\frac{1}{2} (\nabla\phi)^2 + \frac{m^2}{2} \phi^2 + \frac{g}{4!} \phi^4 \right) d^4x.$$

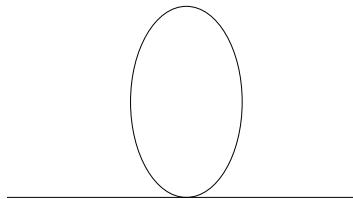
As before, we want to study the Schwinger function $\langle \phi(x_1) \dots \phi(x_N) \rangle$ (it is enough to consider only even N , as the Schwinger function vanishes for odd N). As usual, it is more convenient to consider its Fourier transform. This Fourier transform has the form $G_N(k_1, \dots, k_N) \delta(k_1 + \dots + k_N)$, where G_N is a function on the hyperplane $k_1 + \dots + k_N = 0$.

Consider the Feynman diagram expansion of the function G_N . In ϕ^4 theory we have to sum over graphs whose internal vertices have 4 edges. As usual, we can restrict ourselves to connected graphs, as the sum over disconnected graphs expresses via Schwinger functions with fewer insertion points. In fact, as shown in Kazhdan’s lectures, we can always restrict to 1-particle irreducible graphs (i.e. those which cannot be split by cutting one edge). So from now on we only consider connected, 1-particle irreducible graphs.

From formula (2.4) we get that the superficial divergence index of any graph equals $4 - E$, where E is the number of external edges. Thus, any graph with 2 external edges is quadratically divergent, any graph with 4 external edges is logarithmically divergent, and any graph with > 4 external edges is superficially convergent.

Now we will explain how to renormalize all N -point functions at all orders in g .

First of all, we can exclude graphs with 2 external edges which connect to the same internal vertex, e.g.



and all those which contain such a loop-like graph as a subgraph. Indeed, these graphs produce a constant function of k^2 , so they can be removed by renormalization of mass. More precisely, there exists a function $P(\Lambda, g) = gP_1(\Lambda) + g^2P_2(\Lambda) + \dots$, such that the sum of terms corresponding to all graphs for the theory with mass m^2 , computed with the cutoff propagator (see lecture 1) equals the sum of terms corresponding only to graphs without loop-like subgraphs, but with mass $M(\Lambda)$ such that $M^2 = m^2 + P(\Lambda, g)$. So we can assume from the beginning that we have a theory with mass $M(\Lambda)$ and not worry about loop-like graphs.

Now we have no divergent graphs with one internal vertex, so we have no corrections to make in the first order in g . Let us look at the second order in g . In this case we have the following bad graphs:



Denote the first graph by Γ_2 and the second by Γ_4 .

Thus, in order g^2 we have a problem in the 2-point and the 4-point functions. The problem in the 2-point function is created by Γ_2 . The graph Γ_2 diverges quadratically. Therefore, the term corresponding to Γ_2 computed using the cutoff propagator is a function of k^2, Λ of the form $g^2\Sigma_2(k^2, \Lambda)$, where

$$(2.15) \quad \Sigma_2(k^2, \Lambda) = Ak^2 \ln(\Lambda/\mu) + B(\Lambda) + O(1), B(\Lambda) = B_0\Lambda^2(1 + o(1)), \Lambda \rightarrow +\infty$$

This asymptotics follows from the fact that the first derivative of (2.15) with respect to k^2 is logarithmically divergent, and the second one is convergent.

Another problem we have is in the 4-point function, created by the graph Γ_4 . This graph is logarithmically divergent. This means, if we compute the term corresponding to this graph using the cutoff propagator, we will get a function of k_i, Λ of the form $g^2\Theta_2(k_1, k_2, k_3, \Lambda)$, where

$$(2.16) \quad \Theta_2(k_1, k_2, k_3, \Lambda) = C \ln(\Lambda/\mu) + O(1), \Lambda \rightarrow +\infty$$

This asymptotics follows the fact that the derivative of the integral corresponding to Γ_4 with respect to k_i is convergent.

In order to renormalize the 2- and 4- point functions in order g^2 , we choose renormalized functions $\Sigma_2^R(k^2), \Theta_2^R(k_1, k_2, k_3)$. Here Σ_2^R is an arbitrary function of k^2 whose second derivative is given by the (convergent!) integral obtained by applying $\left(\frac{d}{dk^2}\right)^2$ to the quadratically divergent integral corresponding to Γ_2 . Analogously, Θ_2^R is an arbitrary function of k_1, k_2, k_3 whose derivatives are given by the convergent integrals obtained by differentiating the logarithmically divergent integral corresponding to Γ_4 . It is clear that the function Σ_2^R is defined uniquely up to addition of a function of k^2 of the form $ak^2 + b$, and the function Θ_2^R is defined uniquely up to addition of a constant.

Now we will make second order corrections to the coefficients of $(\nabla\phi)^2, \phi^2$ and ϕ^4 in the Lagrangian. Namely, we will consider a new Lagrangian of the form

$$(2.17) \quad \mathcal{L} = \int \left(\frac{1}{2}(\nabla\phi)^2 + \frac{M^2}{2}\phi^2 + \frac{g}{4!}\phi^4 + \frac{\alpha(\Lambda)}{2}(\nabla\phi)^2 + \frac{\beta(\Lambda)}{2}\phi^2 + \frac{\gamma(\Lambda)}{4!}\phi^4 \right) d^n x,$$

Let $\Sigma'_2(k^2, \Lambda, \alpha, \beta, \gamma)$, $\Theta'_2(k_1, k_2, k_3, \Lambda, \alpha, \beta, \gamma)$ be the functions Σ_2 , Θ_2 for this Lagrangian and the cutoff propagator. We will choose the functions α, β, γ in such a way that

$$(2.18) \quad \begin{aligned} \lim_{\Lambda \rightarrow \infty} \Sigma'_2(k^2, \Lambda, m, \alpha(\Lambda), \beta(\Lambda), \gamma(\Lambda)) &= \Sigma_2^R(k^2), \\ \lim_{\Lambda \rightarrow \infty} \Theta'_2(k_1, k_2, k_3, \Lambda, m, \alpha(\Lambda), \beta(\Lambda), \gamma(\Lambda)) &= \Theta_2^R(k_1, k_2, k_3), \end{aligned}$$

As we are working modulo g^3 , we can choose α, β, γ in the form $\alpha = g^2\alpha_2$, $\beta = g^2\beta_2$, $\gamma = g^2\gamma_2$, where $\alpha_2, \beta_2, \gamma_2$ are independent of g . It is easy to check that in order for (2.18) to hold, the functions $\alpha_2, \beta_2, \gamma_2$ should have the following asymptotics:

$$(2.19) \quad \begin{aligned} \alpha_2 &= g^2(A \ln(\Lambda/m) + D_1) + o(1), \beta_2 = g^2(-B(\Lambda) + D_2) + o(1), \\ \gamma_2 &= g^2(C \ln(\Lambda/m) + D) + o(1), \Lambda \rightarrow +\infty, \end{aligned}$$

where D_1, D_2 depend on the choice of Σ_2^R , and D depends on the choice of Θ_2^R . Of course, there are many ways to choose such functions, but they are unique up to adding terms $o(1)$, $\Lambda \rightarrow \infty$.

Thus, we have renormalized the graphs Γ_2, Γ_4 . This removes divergence in all correlation functions modulo g^3 . Thus, all correlation functions of our theory are now defined modulo g^3 .

Now we proceed inductively in the order of g . Suppose we have removed divergences and defined all correlation functions modulo g^K . Consider the $2N$ -point function (for the deformed Lagrangian and the cutoff propagator) modulo g^{K+1} :

$$(2.20) \quad F_{2N} = \sum_{j=0}^{K-1} g^j F_{2N,j}^R + g^K F_{2N,K}$$

(the superscript R means that the corresponding coefficient has already been renormalized). The term $F_{2N,K}$ is represented by the sum over all graphs with K internal vertices. This sum has superficial divergence index $4 - 2N$. Therefore, the second derivative of F_2 by k^2 , the first partial derivatives of F_4 , and F_{2N} , $N \geq 3$, are superficially convergent. The crucial fact for renormalization theory, which follows from the ‘‘Strong Weinberg Theorem’’ (see Lecture 1), is

Proposition 2.2 *There exists finite limits, as $\Lambda \rightarrow \infty$, of the functions $F_{2N,K}(k_1, \dots, k_{2N-1}, \Lambda)$, $\nabla_k F_{4,K}(k_1, k_2, k_3, \Lambda)$, and $\left(\frac{d}{dk^2}\right)^2 F_{2,K}(k^2, \Lambda)$ $N \geq 3$.*

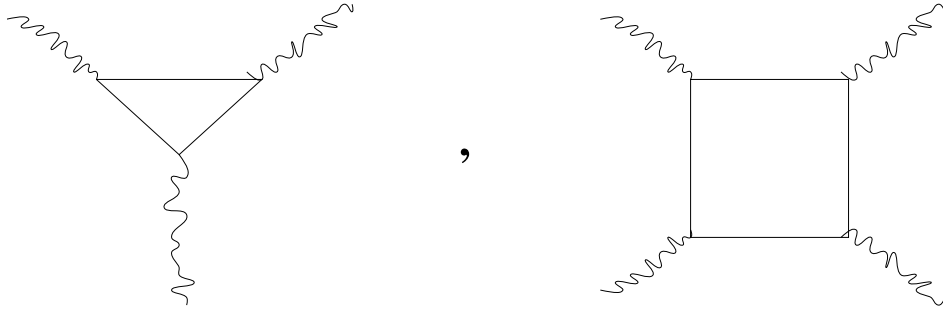
Remark. For ϕ^4 theory, this proposition holds for the term corresponding to each particular graph, but in general (for example, for theories with gauge fields) this is not the case: the sum over all graphs may have a meaning while each individual graph does not. However, an analogue of Proposition 2.2 (for the sum over all graphs) holds in any renormalizable theory.

Proposition 2.2 allows us to fulfil the induction step. It shows that the function F_2 (in the limit) is defined up to adding $ak^2 + b$, the function F_4 is defined up to adding a constant, and F_{2N} , $N \geq 3$, is defined uniquely. So one can choose renormalized functions $F_{2N,K}^R$ and make corrections in the Lagrangian, $\alpha \rightarrow \alpha + g^K\alpha_K$, $\beta \rightarrow \beta + g^K\beta_K$, $\gamma \rightarrow \gamma + g^K\gamma_K$, to compensate the divergence in $F_{2N,K}$ and obtain $F_{2N,K}^R$ instead of it. This procedure is completely analogous to the one for order g^2 . In this way we will complete the renormalization in order K .

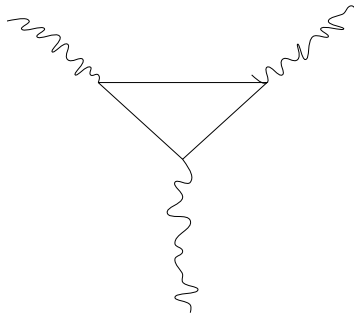
Remark 1. At every step of our renormalization procedure we had to choose 3 constants of integration. This may create an impression that we get a family of theories parametrized by 3 infinite sequences of constants. However, it is easy to see that in fact we get a family of theories parametrized by only 3 constants. This means that any 4 invariants attached to the theory (for example, values of the 2-point function at 4 points in spacetime) are linked by a universal functional relation.

Remark 2. Even if in the original theory certain critical or subcritical interactions were not present, they may appear in the process of renormalization. In general, renormalization brings in all missing critical and subcritical terms, unless there is a symmetry which prevents it from doing so. Let us demonstrate it by a few examples.

Example 1. In the process of renormalization of the Lagrangian (2.12) in 4 dimensions we will be forced to introduce the subcritical term ϕ^3 and the critical term ϕ^4 , in order to remove logarithmic divergence in the graphs



where wavy lines correspond to bosons and straight ones to fermions. However, the critical term $\phi^2 \nabla \phi$ will not appear, since there is no Poincare invariant expression of this form. In terms of graphs, this means that the graph



whose superficial divergence is linear, in fact diverges only logarithmically, because of cancellations in the integrand caused by Poincare symmetry; so there is no linear divergence to compensate and hence no need for $\phi^2 \nabla \phi$ to appear.

Example 2. In ϕ^4 theory in 4 dimensions, the subcritical term ϕ^3 does not appear in renormalization, since it is not preserved by the symmetry $\phi \rightarrow -\phi$ of the original theory. In the language of graphs, this is clear: there is no graphs with 3 external edges, so there is no divergence to compensate by ϕ^3 .

In a general field theory, every type (in terms of external edges) of a superficially divergent graph corresponds to a number of critical and subcritical terms in the Lagrangian, which should be renormalized in order to compensate the divergence in the corresponding graph. More precisely, divergent terms which are quadratic in k correspond to terms in the Lagrangian which have two derivatives by x , linear terms in k correspond to terms with one derivative, and constant divergencies correspond to terms without derivative.

Remark. It follows from formula (2.4) that in a renormalizable theory, all divergences in $N \geq 2$ -point correlation function are no worse than quadratic. So the coefficient of k^2 in the 2-point function if divergent at most logarithmically. Therefore, if the quadratic forms Q_i have to be renormalized (like in ϕ^4 theory), they will be multiplied by coefficients which depend on the cutoff parameter Λ at worst logarithmically. This shows that the dimension of Q_i and hence of ϕ_i survives renormalization to all finite orders in the asymptotic expansion.

Lecture 3: Perturbative renormalization (continued)

Edward Witten³

October 1996

In this lecture we will discuss composite operators, their renormalization and operator product expansion (OPE).

3.1. Local functionals in a classical field theory.

Consider an n -dimensional classical field theory with spacetime V and a Lagrangian \mathcal{L} of the form (2.3), with fields ϕ_1, \dots, ϕ_N . Let X be the space of classical solutions for this theory. We want to consider functions on X called local functionals, which are defined as follows.

Let $x \in V$ be a point in the spacetime.

Definition A local functional at x is a function of the fields and finitely many derivatives of the fields, evaluated at x .

Example In the theory of a scalar bosonic field,

$$\phi^1(x), \phi'(x)^2, \phi^2(x)\phi'(x)^2$$

are local functionals, but $\phi(x) + \phi(2x)$ is not.

From the previous lecture it is clear how to define dimension of a local functional. We want to consider only homogeneous functionals of finite dimension and their finite linear combinations. Therefore, if a field has positive dimension (which is always the case if $n > 2$), we only consider polynomial functionals of this field. However, in 2-dimensional theories, where bosonic fields are dimensionless, it is reasonable and useful to consider more general functions of them (for example, the Lagrangians of Toda theories contain expressions of the form e^ϕ).

If $n > 2$, the space of all functionals of a given dimension is finite-dimensional, but if $n = 2$, it is infinite-dimensional.

Since elements of X satisfy the classical field equations, the same local functional can be written in different ways. For example, in ϕ^4 -theory the classical field equation is

$$(3.1) \quad \Delta\phi = m^2\phi + \frac{g}{3!}\phi^3,$$

so the left and the right hand sides of (3.1) evaluated at $x \in V$ represent the same local functionals of ϕ . Thus, local functionals are all possible differential expressions in ϕ modulo the classical field equations.

Remark 1. One should be careful to distinguish between two different notions of field dimension which arise in field theory. The first is “the engineering dimension” and says in what units the field is measured (if the units are cm^{-d} then the engineering dimension is d). The second is “the scaling dimension”, which is the dimension we have been talking about. These two dimensions are not always the same. For example, in the theory of one field ϕ defined by (3.1) the engineering dimension of the field $m^2\phi$ is 3 (it is measured in cm^{-3}), while the scaling dimension is the same as that of ϕ , i.e. 1.

It is the basic principle of physics that all meaningful expressions and equations are homogeneous with respect to the engineering dimension. In particular, it is true for the field equations in field theory (e.g. (3.1)), which implies that engineering dimension defines a grading on the space of local functionals. On the other hand, whenever a field theory is not scale invariant, its field equations are not

³Notes by Pasha Etingof and David Kazhdan, TeXnical editing Misha Verbitsky

homogeneous with respect to the scaling dimension, so scaling dimension defines a filtration, rather than a grading, on the space of local functionals. This means, the space of all local functionals is a union of subspaces of functionals of dimension $\leq d$, and a functional is said to be of dimension d if it is of dimension $\leq d$ but not $\leq d - 1$. On the other hand, if the theory happens to be scale invariant, scaling dimension does define a filtration on the space of local functionals.

In this lecture, we will not use engineering dimension, and the word “dimension” will always mean scaling dimension.

Remark 2. From a local functional \mathcal{O} one can obtain more general functionals on X of the form $\int \mathcal{O}(\phi, x) d\mu(x)$, where $d\mu(x)$ is a (generalized) density on V . Using this operation, one can obtain from any local functional all derivatives of this functional (it is enough to take for $d\mu(x)$ all possible densities supported at x). Therefore, without loss of information we could consider local functionals modulo the image of derivatives. However, for the purposes of this lecture this is not necessary.

3.2. Quantization of local functionals in a free theory

We have seen (see Bernstein’s Lectures and Witten’s problem sets) that the space X of classical solutions of a meaningful classical field theory always carries a natural closed 2-form Ω . If this form is nondegenerate, X is a symplectic manifold. In this case, suitable functions on X form a Poisson algebra.

In the quantum theory, the space X should be quantized, and the Poisson algebra of functions on X should become the algebra of operators (observables) in some Hilbert space of states. In particular, we should be able to assign an operator to every local functional.

If V is a Minkowski space, and the field theory satisfies the Wightman axioms (for example, the free theory), the Hilbert space of states \mathcal{H} is constructed as described in Kazhdan’s lectures. In this case, Wightman fields $\phi(x)$ are distributions on V with values in the space of operators on the subspace $\mathcal{D} \subset \mathcal{H}$ of muptiparticle states. This means, for any Schwarz function f on V there is an honest operator $\varphi(f)$ on \mathcal{D} .

However, we would like to have more general operators of the form $\phi^N(x)$, $(\nabla\phi)^2(x)$, etc. which correspond to local functionals in the classical theory. That is, for any Schwarz function f we want to have operators $\varphi^N(f)$, $(\nabla\varphi)^2(f)$, etc.

Unfortunately, such operators are not automatically defined. For example, even $\varphi^2(f)$ does not, in general, make sense. Indeed, let $\delta_\varepsilon = e^{-|x|^2/\varepsilon} (2\pi/\varepsilon)^{n/2}$ be the smooth approximation to the δ -function, and let $\varphi \otimes \varphi$ be the operator-valued distribution on $V \times V$ given by $\varphi \otimes \varphi(f_1 \otimes f_2) = \varphi(f_1)\varphi(f_2)$. The natural definition of $\varphi^2(f)$ would be that $\varphi^2(f)$ is the limit, as $\varepsilon \rightarrow 0$, of the operator $\varphi_\varepsilon^2(f) := \varphi \otimes \varphi(f(x)\delta_\varepsilon(x-y))$ (in the sense of convergence of matrix elements). However, it is easy to see that this limit does not usually exist, as Wightman functions usually have singularities on the diagonals.

One way to deal with this problem is to say that an “operator” A is just a collection of its matrix elements, i.e. a collection of correlation functions $\langle \phi(y_1) \dots \phi(y_l) | A | \phi(z_1) \dots \phi(z_r) \rangle$. If we accept this point of view, we might as well forget about the Hilbert space of states, i.e. perform a Wick rotation from Minkowski space to Euclidean space, and consider Schwinger functions instead of Wightman functions.

So from now on we will consider only the Euclidean situation, in which we will mean by an operator A a collection of functions

$$\langle \phi(y_1) \dots \phi(y_l) | A | \phi(z_1) \dots \phi(z_r) \rangle$$

with certain properties. Roughly, an operator is just a symbol which can be inserted in a correlation function.

Remark. Of course we should remember at all times that A is not really an operator and does not act

in any Hilbert space. However, the information we get from studying A in the Euclidean situation can be used (after Wick rotation) for studying the Minkowski situation.

Consider now the problem of quantization of local functionals. By the definition, in order to quantize the local functional $\mathcal{O}(x)$, we should say what the functions $\langle \phi(y_1)\dots\phi(y_l)\mathcal{O}(x)|\phi(z_1)\dots\phi(z_r)\rangle$ are. As our theory is Euclidean, the order of factors in the correlation function does not matter (any two distinct points are space-like separated). So, in order to define $\mathcal{O}(x)$, it is enough to define the correlation functions $\langle \phi(y_1)\dots\phi(y_r)\mathcal{O}(x)\rangle$ for all r .

Consider the theory of a scalar bosonic field in $n > 2$ dimensions, with the Lagrangian

$$(3.2) \quad \mathcal{L}(\phi) = \int \left(\frac{1}{2}(\nabla\phi)^2 + \frac{m^2}{2}\phi^2 + Q(\phi) \right) d^n x,$$

We will give all definitions and constructions for this example. In other field theories, everything is done in a similar manner.

We first consider the case when the theory is free ($Q = 0$). In a free theory, quantization is done with the help of normal ordering, as follows. Consider a local functional \mathcal{O} of degree J in ϕ . According to Feynman rules, in order to compute the correlation function $\langle \phi(y_1)\dots\phi(y_r)\mathcal{O}(x)\rangle$, we should consider all graphs with r external vertices and only one internal vertex v , which has J edges. We should put a certain function at the vertex v , and compute the term (amplitude) corresponding to this graph as usual in Feynman calculus. If $J > 1$, we will run into trouble: we will get some graphs with loops going from v to itself, and the integration over the loops is divergent. The easiest remedy is to ignore all such graphs. Then we obtain certain correlation functions, which define some operator. This operator is denoted by $:\mathcal{O}(x):$ and called the normal ordering of $\mathcal{O}(x)$. Apriori, $\mathcal{O}(x)$ does not make sense as an operator, while $:\mathcal{O}(x):$ does. We call operators of the form $:\mathcal{O}(x):$ composite operators. When no confusion is possible, we will drop the dots and write $\mathcal{O}(x)$.

Thus, we have assigned canonically to each differential polynomial \mathcal{O} in ϕ an operator $:\mathcal{O}:$ in the free theory. However, recall that two different polynomials $\mathcal{O}_1, \mathcal{O}_2$ might define the same local functional. So does the map $\mathcal{O} \rightarrow :\mathcal{O}:$ actually define canonically a quantization of the local functional represented by \mathcal{O} ? The answer turns out to be positive: whenever polynomials $\mathcal{O}_1, \mathcal{O}_2$ define the same local functional, the operators $:\mathcal{O}_1:; :\mathcal{O}_2:$ coincide. To check this is an easy exercise. For example, the polynomials $\mathcal{O}_1 = \phi\Delta\phi, \mathcal{O}_2 = m^2\phi^2$ define the same local functional, and the operators $:\phi\Delta\phi:; :m^2\phi^2:$ coincide.

3.3. Multiplication of composite operators.

It is easy to multiply composite operators supported at different points. That means, given two composite operators $\mathcal{O}, \mathcal{O}'$, and points $x, x' \in V, x \neq x'$, it is easy to define an operator $\mathcal{O}(x)\mathcal{O}'(x')$, using the Feynman diagram expansion. In this case we will have two special vertices v, v' , and we should sum over all ways of connecting them which involve no loops from a vertex to itself. However, multiplication of two composite operators $\mathcal{O}, \mathcal{O}'$ at the same point x cannot be defined in the same way: on each edge connecting v with v' , we will have to evaluate the Green's function at (x, x) , where it is singular.

To avoid this problem, one can try to smear the operators, i.e. consider the operator $\mathcal{O}(f) = \int f(x)\mathcal{O}(x)d^n x$, where f is a compactly supported smooth function on V . Such smeared operators can sometimes be multiplied: for example, we can define the product $\phi(f_1)\dots\phi(f_s)$. (see Kazhdan lectures).

However, most composite operators cannot be multiplied, even after smearing. Indeed, it can be deduced from reflection positivity that for any complex operator \mathcal{O} $\langle \mathcal{O}(x)\mathcal{O}^*(x')\rangle \sim |x - x'|^{-2[O]}$, $x \rightarrow x'$ where $[O]$ is the dimension of $[O]$, and \mathcal{O}^* is the conjugate of \mathcal{O} . On the other hand, for a real operator \mathcal{O} , in order for the product $\mathcal{O}(f)\mathcal{O}(f')$ to be defined, it is necessary and sufficient that the functions $\langle \phi(y_1)\dots\phi(y_r)\mathcal{O}(x)\mathcal{O}(x')\rangle$ be distributions. The function $|x|^{-l}$ defines a distribution (without

regularization) if and only if it is locally L^1 , i.e. iff $l < n$. Thus, the product $O(f)O(f')$ is automatically defined iff $[O] < n/2$.

Remark. If the Minkowski situation, the product $O(f)O(f')$ exists (in the sense of correlation functions) if and only if the “operator” $O(x)$ is an operator-valued distribution, so that $O(f)$ is actually an honest operator (i.e. matrix elements of $O(f)$ are matrix elements of some operator on \mathcal{D}). Indeed, the “if” part is clear. The “only if” part: for any f_1, \dots, f_N we can regard the product $X = O(f)\varphi(f_1)\dots\varphi(f_N)$ as a linear functional on \mathcal{D} . The fact that $O(f)^2$ is defined implies that this functional has finite L^2 -norm, as $\|X\|^2 = \langle \varphi(f_N)\dots\varphi(f_1)O(f)^2\varphi(f_1)\dots\varphi(f_N) \rangle$. Therefore, X corresponds to a vector in \mathcal{H} , Q.E.D.

3.4. Operator product expansion (OPE) in the free theory

Composite operators in a reasonable field theory have an important algebraic property, which is called the operator product expansion (OPE). This is a property of algebraic nature, saying roughly that composite operators form an algebra of a certain kind. It does not follow obviously from Wightman axioms, but on the other hand is often satisfied when Wightman axioms fail. This property is actually useful in practice.

We will now define and compute the OPE for the free theory of a scalar bosonic field. The idea is to formalize the behavior of the product $O(x)O'(x')$ as $x \rightarrow x'$. In the classical theory, this is accomplished by the usual Taylor series. Indeed, we have

$$(3.3) \quad O(x)O'(x') = O(x')O'(x') + (x - x')_\mu \partial_\mu O(x')O'(x') + \dots$$

Thus, for any two local functionals O, O' the product $O(x)O'(x')$ decomposes in a series in products of powers of $(x - x')_\mu$ whose coefficients are other local functionals.

In the quantum theory, the product $:O(x)::O'(x'):$, as we have seen, is singular at $x = x'$. Thus, we have to compute the asymptotic expansion of this product near $x = x'$.

We will assume that O, O' are represented by some monomials in ϕ and its derivatives. According to Feynman rules, we should assign to both O, O' a vertex, say v and v' . The numbers of edges at v, v' are J, J' , where J, J' are the degrees of O, O' as monomials of ϕ . At both v, v' , we have to place symmetric distributions in J, J' variables corresponding to O, O' . Now we have to sum amplitudes over all ways of connecting some edges going from v with some edges going from v' . The amplitude of each particular graph is computed as follows: at each edge going from v to v' we put the function $D(y - z)$, the two-point Schwinger function of the free theory (the Green's function for the Helmholtz equation). On each remaining external edge we put $\phi(y)$. Then we compute formally the product of things on vertices and edges, and put it under the normal ordering sign. Since for each number K of connecting edges, we have only one possible graph, we should just compute the amplitude of this graph, and then multiply it by the number of ways to get this graph, i.e. by $J!J'/K!(J - K)!(J' - K)!$.

This will give us a formula of the form

$$(3.4) \quad :O(x)::O'(x'):= \sum_{k=1}^L : \tilde{O}_k(x)\tilde{O}'_k(x') : E_k(x - x'),$$

where $:\tilde{O}_k::, :\tilde{O}'_k:$ are some local functionals, and $E_k(y)$ are polynomial differential operators in $D(y)$. Here $:O(x)O'(x'):$ is the amplitude of the graph in which the vertices v and v' are not connected.

Now we can use the Taylor formula (3.3) inside of the normal ordering, to get the following (already infinite) expansion:

$$(3.5) \quad :O(x)::O'(x'):= \sum_s :O_s(x') : D_s(x - x'),$$

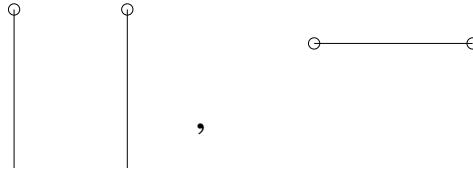
where O_s is local functional, and $D_s(y)$ is a product of the function $E_k(y)$ for some k and a polynomial in y .

Definition. Expansion (3.5) is called the operator product expansion (OPE).

Now recall that the Green's function $D(x)$ has the following behavior at $x = 0$: $D(x) \sim |x|^{2-n}$. Therefore, all functions $d_k(x)$ have the property $|d_k(x)| \leq C_k|x|^{r_k}$ for small x , where the numbers r_k are almost all positive. This proves an important property of the OPE: only finitely many terms are singular. However, these singular terms carry the most interesting information in many situations.

Remark. We see that the operator product expansion (3.5) is in fact the quantum analogue of the Taylor expansion (3.3). Indeed, it is an expansion of the form similar to (3.3), but also involving finitely many singular terms at $x = x'$.

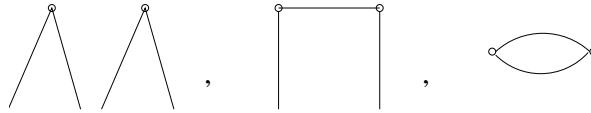
Let us now consider examples of OPE (we will always write the finite formula (3.4)). Let us first compute $\phi(x)\phi(x')$. Using the above rules, we get two graphs:



which yields the formula

$$(3.6) \quad \phi(x)\phi(x') =: \phi(x)\phi(x') : + D(x - x').$$

Now we compute $: \phi^2(x)\phi^2(x') :$. This product gives us three graphs:



so we get the formula

$$(3.7) \quad : \phi^2(x) : : \phi^2(x') : =: \phi^2(x)\phi^2(x') : + 4 : \phi(x)\phi(x') : D(x - x') + 2D^2(x - x').$$

In general, it is easy to prove that

$$(3.8) \quad : \phi^J(x) : : \phi^{J'}(x') : = \sum_{K=0}^{\min(J, J')} \frac{J!J'!}{K!(J-K)!(J'-K)!} \times D(x - x')^K : \phi^{J-K}(x)\phi^{J'-K'}(x') :$$

The last formula can be written more nicely using generating functions:

$$(3.9) \quad : e^{\alpha\phi(x)} : : e^{\beta\phi(x')} : = e^{\alpha\beta D(x-x')} : e^{\alpha\phi(x)} e^{\beta\phi(x')} :,$$

where α, β are constants.

If we take α, β to be any differential operators on V with constant coefficients, this formula remains valid. (α acts on x , β on x'). In this form, formula (3.7) represents the most general OPE for the free theory of a scalar bosonic field.

Remark. Dimensions of fields in quantum theory may differ from the dimensions of their classical analogues. For example, in a (2-dimensional) classical field theory, if a field ϕ is 0-dimensional,

so is $f(\phi)$, where $f : \mathbb{R} \rightarrow \mathbb{R}$ is any function. However, in quantum theory a renormalized composite operator $f(\phi)$ may acquire a nontrivial (anomalous) dimension. For example, in the free 2-dimensional theory of a massless scalar bosonic field, $D(y) = -\ln|y|$, so we have from (3.9):

$$\langle : e^{i\alpha\phi(x)} :: e^{-i\alpha\phi(y)} : \rangle = |x - y|^{-\alpha^2}.$$

But we know that $\mathcal{O}(x)\mathcal{O}^*(y) \sim |x - y|^{-2[\mathcal{O}]}$. Thus, the scaling dimension of $e^{i\alpha\phi}$ is $d(\alpha) = \alpha^2/2$. Observe that the function d is not linear in α , so for quantum dimensions, to the contrary with the classical dimensions, $[:\mathcal{O}_1\mathcal{O}_2:] \neq[:\mathcal{O}_1:] +[:\mathcal{O}_2:]$.

The dimension function $d(\alpha) = \alpha^2/2$ appears in the theory of vertex operator algebras, namely, in the Frenkel-Kac vertex operator construction.

In an interacting theory, even polynomial fields can have different dimensions quantum-mechanically than they do classically. This is essential, however, mostly in the non-perturbative setting. In the perturbative setting we can count dimensions as in the classical theory, since dimensional anomalies are infinitesimally small.

3.5. Normal ordering and renormalization.

Now we will reformulate the definition of normal ordering in the free theory in terms of renormalization theory. This reformulation will be crucial in understanding what is the analogue of normal ordering in interacting theories.

At the beginning of this lecture, we defined normal ordering of a local functional by formally throwing away graphs with self-loops, which produced divergence. Another way of dealing with self-loops is renormalizing them, as we did in two previous lectures. Suppose we want to renormalize all local functionals of dimension $\leq d$. For this purpose we replace the standard propagator $\frac{1}{k^2+m^2}$ with a cutoff propagator $\frac{1}{k^2+m^2} \left(\frac{\Lambda^2}{k^2+\Lambda^2} \right)^l$, where l is sufficiently big. Now all the loop integrals converge, and for any local functional \mathcal{O} of degree $\leq d$ we can consider the cutoff correlation function $\langle \phi(y_1)\dots\phi(y_r)\mathcal{O}(x) \rangle_\Lambda$. If we take $\Lambda \rightarrow \infty$, we will of course find that the limit does not exist. However, it is not difficult to prove the following.

Let A_d be the space of local functionals of dimension $\leq d$.

Proposition 3.1. There exists a Λ -dependent linear map $R_\Lambda : A_d \rightarrow A_d$, strictly triangular with respect to the filtration of A_d by dimension, such that for any local functional $\mathcal{O} \in A_d$ there exists a limit

$$(3.10) \quad \lim_{\Lambda \rightarrow \infty} \langle \phi(y_1)\dots\phi(y_r)R_\Lambda\mathcal{O}(x) \rangle_\Lambda,$$

Thus, Proposition 3.1 allows us to assign to every local functional \mathcal{O} an operator $\tilde{\mathcal{O}}$, which is, by definition, the operator whose matrix elements are the limits of the corresponding matrix elements of $R_\Lambda\mathcal{O}$. However, given \mathcal{O} , the operator $\tilde{\mathcal{O}}$ is defined a priori noncanonically, as the map R_Λ is not unique: it is defined up to left multiplication with a Λ -independent strictly triangular map.

That is, $\tilde{\mathcal{O}}$ is defined uniquely up to adding composite operators of lower dimension. This shows that the space of composite operators is naturally a filtered object (by dimension), and not a graded object.

Remark. Of course, if the theory is not classically scale invariant, we saw that there is no grading on functionals already at the classical level. The statement here is that even for a classically scale-invariant theory, where the space of classical functionals automatically has a grading by scaling dimension, the grading is usually lost in the process of quantization. An exception is a free scale-invariant theory, where there is a canonical quantization by normal ordering, and therefore the grading survives quantization.

Thus, in general we may be able to quantize naturally the space of local functionals of dimension $\leq d$, but not every functional separately.

3.6. Composite operators in an interacting critical theory.

Now consider an interacting renormalizable field theory. As a model example we consider the Lagrangian (3.2) and take $Q(\phi) = \frac{g}{4!}\phi^4$. Let O be any local functional represented by a monomial. To quantize O , we proceed as in the free case, but we will formulate everything in a slightly different language.

We want to consider correlation functions $\langle \phi(y_1)\dots\phi(y_r)O(z) \rangle$. These correlation functions can be viewed as the ε -coefficient in the usual Schwinger functions for a perturbed Lagrangian, of the form

$$(3.11) \quad \mathcal{L}_\varepsilon(\phi) = \mathcal{L}(\phi) + \varepsilon O(\phi, z),$$

where ε is a formal variable such that $\varepsilon^2 = 0$. In the language of Feynman diagrams, this means that we are introducing an additional vertex v corresponding to O , and summing over all graphs which contain exactly one such vertex (with no self-loops at this vertex) and are otherwise as usual.

Remark. One should remember that there is no momentum conservation at the vertex v .

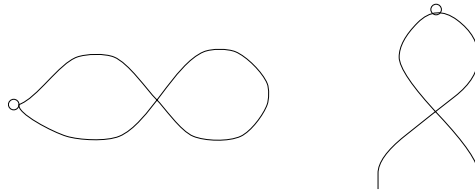
In general, such an alteration will worsen the divergence properties of the Feynman graphs. More precisely, now the superficial divergence of a graph with E external edges is given by $div(\Gamma) = [O] - E$. However, we can renormalize these divergences, using the cutoff propagator considered in the previous section. Then, analogously to Proposition 3.1, one can prove the following.

Proposition 3.2. There exists a Λ -dependent linear map $R_\Lambda : A_d \rightarrow A_d$, triangular (in general, not strictly) with respect to the filtration of A_d by dimension, such that for any local functional $O \in A_d$ there exists a limit

$$(3.12) \quad \lim_{\Lambda \rightarrow \infty} \langle \phi(y_1)\dots\phi(y_r)R_\Lambda O(x) \rangle_\Lambda,$$

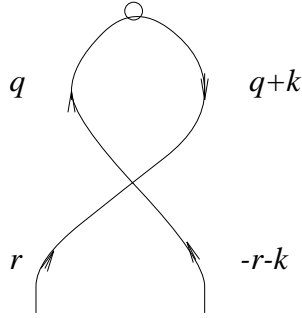
As in the free theory, this Proposition allows to quantize the space of local functionals of dimension $\leq d$, but in general there is no canonical quantum analogue for each classical local functional. This non-uniqueness is not only due to the non-uniqueness of renormalization, but also due to the non-uniqueness of representation of a given functional by a differential polynomial.

Example. Let $O = \phi^2/2$. Let us compute the renormalization of O of order g . The only divergent graphs we have in this order are



Let us call the first graph by Γ_0 and the second by Γ_2 . The graph Γ_0 is quadratically divergent. If we replace the usual propagator with the cutoff propagator, the integral will converge to a Λ -dependent constant of the form gC_Λ , where C_Λ , which grows quadratically in Λ .

Now consider the graph Γ_2 . As usual, it is more convenient to work in the momentum space, i.e. consider the Fourier transform of the term corresponding to this graph. Let k_1, k_2, k be the corresponding momentum variables. Then the Fourier transform of the term corresponding to Γ_2 is of the form $F_2\delta(k_1 + k_2 + k)$, where F_2 is a function on the plane $k_1 + k_2 + k = 0$. Set $k_1 = r$, then $k_2 = -r - k$ (on this plane). Thus, $F_2 = F_2(r, k)$, and its order g correction is the amplitude of the Feynman diagram



This amplitude is given by the integral

$$(3.13) \quad I(k) = -\frac{g}{2} \int \frac{d^4 q}{(2\pi)^4} \frac{1}{(q^2 + m^2)((q+k)^2 + m^2)}.$$

This integral is logarithmically divergent. If we replace the usual propagator with the cutoff propagator, the integral will converge to a function $I_\Lambda(k)$, which has the asymptotic behavior

$$I_\Lambda(k) = -\frac{gA}{2} \ln(\Lambda/\mu) + O(1), \Lambda \rightarrow \infty$$

where A is a constant.

This shows that if instead of $\phi^2/2$ we use the renormalized functional

$$(3.14) \quad \left(\frac{\phi^2}{2}\right)_\Lambda := \frac{\phi^2}{2} \left(1 + \frac{gA}{2} \ln(\Lambda/\mu)\right) - gC_\Lambda,$$

the matrix elements will have a finite limit modulo g^2 . Observe that the constant μ can be chosen arbitrarily, so the renormalization is not canonical already in this case.

3.7. Stability of the classical field equations under quantization

In a free theory, we know that the classical field equations are also satisfied quantum-mechanically. This means, if the classical equations are $P\phi = 0$, where P is a linear differential operator, then the correlation functions in the quantum theory satisfy the equation $P_x \langle \phi(y_1) \dots \phi(y_r) \phi(x) \rangle = 0$ outside of the diagonals $x = y_i$. For example, the Green's function $D(x-y) = \langle \phi(x) \phi(y) \rangle$ is the fundamental solution of the equation $Pf = 0$.

In an interacting theory, there are some problems. Namely, since the classical field equations for an interacting theory are nonlinear (e.g. (3.1)), they do not make sense quantum-mechanically in the setting of Wightman axioms. However, in the OPE setting they make sense in a suitable interpretation, and one can show that they are satisfied.

Remark. We will say that the equation $F(\phi) = 0$ is satisfied in quantum theory (where F is a renormalized local functional) if the function $\langle \phi(y_1) \dots \phi(y_r) F(\phi(x)) \rangle$ vanishes outside of the diagonals $x = y_i$.

We will show that the classical field equations are satisfied in ϕ^4 theory. Consider the space of classical local functionals of dimension ≤ 3 , which are Poincare-invariant and odd under the symmetry $\phi \rightarrow -\phi$. This space is 2-dimensional: it has 3 generators ϕ^3 , $\Delta\phi$, and ϕ , but they are linearly dependent, since they satisfy relation (3.1).

Now consider the quantum theory. Consider composite operators of dimension ≤ 3 , with the same invariance properties as above. In this space we have renormalized operators $\Delta\phi$, ϕ , ϕ^3 . Our goal is to show that, like in the classical theory, they are linearly dependent. This is equivalent to the statement that ϕ satisfies the field equation $C\phi^3 = A\Delta\phi + B\phi$.

We will first prove the validity of the field equation for the theory with the cutoff propagator, since in this theory we do not have divergence problems. In the cutoff theory, the classical field equation is

$$(3.15) \quad P_\Lambda \phi = \frac{g}{3!} \phi^3, P_\Lambda = (\Delta - m^2) \left(1 - \frac{\Delta}{\Lambda^2}\right)^l.$$

Proposition 3.3 Equation (3.15) holds in the quantum theory with the cutoff propagator, where ϕ and ϕ^3 are regarded as composite operators.

Proof. Consider the correlation function $\langle \phi(y_1) \dots \phi(y_r) P_\Lambda \phi(x) \rangle$. Consider the graph decomposition of this function. We should consider all possible graphs with r external edges, any number of 4-valent internal vertices, and a special vertex v with 1 outgoing edge. These graphs can be of two kinds: 1) graphs in which v connects to an external vertex; 2) graphs in which v connects to an internal vertex. The sum over graphs of the first type is the corresponding correlation function for the free theory, so it is supported on the union of diagonals $x = y_i$. Thus, outside of the diagonals we have to sum only over graphs of the second kind. But graphs of the second kind with N external edges are in 1-1 correspondence with usual Feynman graphs (without v) with $N+3$ external edges (this correspondence is obtained by biting off the vertex v and the vertex it connects to). This shows that the sum over graphs of the second kind is just $\frac{g}{3!} \langle \phi(y_1) \dots \phi(y_r) \phi^3(x) \rangle$. Thus, the field equation is satisfied, Q.E.D.

Proposition 3.3 shows that for any Λ , the composite operator $\frac{g}{3!} \phi^3$ is linearly dependent of composite operators which are linear in ϕ . This property has to be preserved in the limit $\Lambda \rightarrow \infty$. This shows that in the renormalized ϕ^4 theory we have an equation $C\phi^3 = P\phi$, where P is a linear differential operator with constant coefficients, and C is a constant. Since the r.h.s. of this equation can only have terms of dimension ≤ 3 , the operator P has to be of the form $A\Delta + B$. Thus in the quantum theory we have the equation

$$(3.16) \quad C\phi^3 = A\Delta\phi + B\phi$$

Remark. Of course, the constants A, B, C are not uniquely defined, as they depend on the choice of the renormalization. The statement is only that for any choice of renormalization, some nontrivial equation of the form (3.16) holds. In other words, the statement is that the space of Poincare invariant, odd composite operators of dimension ≤ 3 is 2-dimensional, i.e. has the same dimension as the corresponding space of classical local functionals.

In general, one can consider the space of local functionals of dimension $\leq N$. Denote its dimension by $d(N)$. One can prove the following proposition.

Proposition 3.4 The space of composite operators of dimension $\leq N$ is also equal to $d(N)$.

Thus, all differential equations which are satisfied classically, have quantum analogues.

3.8. Operator product expansion in an interacting theory.

In a free theory, once we defined composite operators, it was no problem to define the product of several of them supported at different points. The same is true in the interacting theory. Namely, if O_1, \dots, O_N are renormalized composite operators, and $x_1, \dots, x_N \in V$ distinct points, then the correlation function $\langle \phi(y_1) \dots \phi(y_r) \Theta(x_1) \dots \Theta(x_N) \rangle$ is just the coefficient of $\varepsilon_1 \dots \varepsilon_N$ in the usual Schwinger function $\langle \phi(y_1) \dots \phi(y_r) \rangle$ for the Lagrangian

$$(3.17) \quad \mathcal{L}_{\varepsilon_1, \dots, \varepsilon_N}(\phi) = \mathcal{L}(\phi) + \sum_{i=1}^N \varepsilon_i O_i(\phi, x_i),$$

where $\varepsilon_i^2 = 0$. This Schwinger function is, by definition, the limit of the corresponding function computed with the cutoff propagator. It is easy to see, by looking at Feynman diagrams, that this limit is always finite.

Remark. Roughly speaking, this means that once we renormalized each of the operators O_i , we automatically renormalized the product

$$O_1(x_1)\dots O_N(x_N), \quad x_i \neq x_j.$$

Thus, now we are in as good shape as we were in the free theory before we defined OPE. It turns out, we can define OPE in the interacting theory as well. Namely, for any two composite operators O, O' and for any $r \in \mathbb{R}$ there exists a finite collection of composite operators O_1, \dots, O_l such that

$$O(x)O(x') = \sum_{s=1}^l O_s(x')D_s(x-x') + O(|x-x'|^r).$$

We will show how to partially compute the OPE on examples.

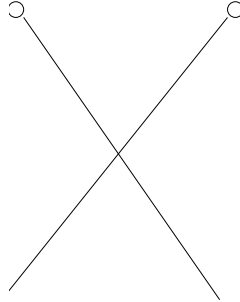
First we consider the product $\phi(x)\phi(x')$, and compute its asymptotics as $x \rightarrow x'$. In this case the expansion is of the form:

$$(3.18) \quad \phi(x)\phi(x') = \phi_R^2(x')f(x-x') + h(x-x') + \text{regular part},$$

where ϕ_R^2 is the renormalized operator ϕ^2 defined by (3.17). (recall that ϕ_R^2 is defined non-uniquely). Let us find functions f, h .

The function $h(z)$ can be found by looking at the 0-point function for the operator $\phi(x)\phi(x')$. Indeed, since $\langle \phi_R^2(x) \rangle = 0$ by the definition, we get $h(x-x') = \langle \phi(x)\phi(x') \rangle$. If we compute the answer modulo g^2 , we get no corrections to the free theory answer, so $h(z) = D(z)$ (see lecture 2).

Now let us find $f(z)$ (we remember that it is defined up to scaling $f \rightarrow a(g)f$, where $a = 1 \pmod{g}$). This function is found from the 2-point function of $\phi(x)\phi(x')$. We have $f(z) = 1 + gf_1(z) + O(g^2)$. The only graph that contributes to $f_1(z)$ is



The amplitude of this graph (in the position space) is

$$(3.19) \quad \begin{aligned} & g \int D(x-z)D(x'-z)D(y_1-z)D(y_2-z)dz \\ &= -\frac{gA}{2} \ln \left| \frac{x-x'}{\mu} \right| D(y_1-x')D(y_2-x') \\ &+ \text{regular part} \end{aligned}$$

(A is the same constant as in (3.17)). This shows that $f_1(z) = -\frac{gA}{2} \ln \left| \frac{x-x'}{\mu} \right|$, and thus

$$(3.20) \quad \begin{aligned} \phi(x)\phi(x') &= \phi_R^2(x') \left(1 - \frac{gA}{2} \ln \left| \frac{x-x'}{\mu} \right| \right) + D(x-x') \\ &+ \text{regular part} + O(g^2). \end{aligned}$$

Now consider a more complicated operator product, for example

$$\phi_R^2(x)\phi_R^2(x').$$

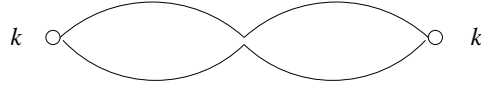
In this case the expansion is of the form

$$(3.21) \quad \begin{aligned} \phi_R^2(x)\phi_R^2(x') &= \phi_R^4(x')f_1(x-x') + (\nabla\phi_R^2)(x')f_2(x-x') \\ &+ \phi_R^2(x')f_3(x-x') + f_4(x-x') \\ &+ \text{regular part,} \end{aligned}$$

where the subscript R means “renormalized composite operator”. Of course the functions f_1, \dots, f_4 will depend on the choice of the renormalization, but they are well defined up to an upper triangular linear transformation.

For the sake of brevity we will only compute the expansion modulo $O(|x-x'|^{-1})$. Thus, the functions f_1, f_2 can be ignored, and we only have to compute f_3, f_4 modulo $O(|x-x'|^{-1})$.

The function f_4 is defined canonically up to multiplication by a scalar. As before, this function is computed using the 0-point function: $f_4(x-x') = \langle \phi_R^2(x)\phi_R^2(x') \rangle$. The only graph that contributes to the 1-st order in g of f_4 is



This shows that

$$f_4(z) = 2D(z)^2 + g\widehat{S(k)} + \text{regular part} + O(g^2),$$

where $S(k)$ is the amplitude (in the momentum space) of the above graph, and hat denotes the Fourier transform. The function $S(k)$ equals $T(k)^2$, where $T(k)$ is the renormalized integral

$$T(k) = C \int_R \frac{dq}{(q^2 + m^2)((k-q)^2 + m^2)},$$

Remark. Since $T(k)$ is defined up to adding a constant, $f_4(z)$ is defined up to adding a multiple of $g\widehat{T(k)}$. Since $\widehat{T(k)}$ is proportional to $D^2(z)$ when $z \neq 0$, adding such multiple is equivalent to multiplying f_4 by $1 + cg$. This freedom is natural, as f_4 is defined up to multiplying by $1 + cg$.

Now let us try to compute $f_3(z)$, which is also defined canonically up to a scalar. For this purpose we should consider the 2-point function of

$$\phi_R^2(x)\phi_R^2(x'),$$

i.e.

$$\langle \phi(y_1)\phi(y_2)\phi_R^2(x)\phi_R^2(x') \rangle.$$

We will work with the Fourier transform of this function, which we denote by $F_2(p_1, p_2, q, q')$ (here p_1, p_2, q, q' are the dual variables to y_1, y_2, x, x').

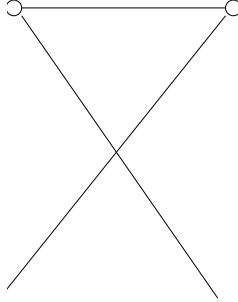
We will try to compute the function $f_3(z)$ by looking at the asymptotic expansion of $F_2(p_1, p_2, q, r-q)$ as $|q| \rightarrow \infty$, and using the following fact from calculus:

Claim Let f be an L^1 -function on an n -dimensional Euclidean space V whose Fourier transform \hat{f} satisfies the inequality $|\hat{f}(q)| \leq C|q|^{-n-N-\varepsilon}$, $\varepsilon > 0$. Then $f \in C^n(V)$.

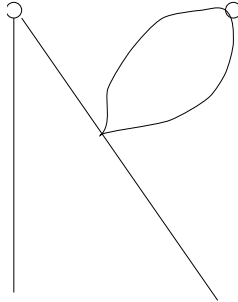
By doing so we will be able to find $f_3(z)$ modulo C^∞ -functions, which is all we want at this point.

We have $f_3(z) = 4D(z) + gf_3^1(z)$. Looking at the Feynman diagram expansion of the 2-point function in the first order of g , we see terms of two kinds: 1) terms where the special vertices v, v' are directly connected with an edge; 2) terms where v, v' are not connected.

The only graph of the first kind that contributes to f_3^1 is



and the only graph of the second kind contributing to f_3^1 is



(this graph occurs 4 times, with 4 different labelings of vertices). It is easy to check directly that both graphs contribute to $\hat{f}_3(q)$ a term of the form $Cq^{-2} \ln|q/\mu| + O(q^{-4+\epsilon})$, where C is a constant. Thus, $f_3^1(z)$ is the Fourier transform of the function $C \ln|q/\mu|/q^2$ modulo $O(|x-x'|)^{-1}$ and $O(g^2)$. As before, the constant μ depends on the choice of renormalization; changing of this constant is equivalent to adding a multiple of $D(z)$ to $f_3^1(z)$, which is the same as multiplying $f_3(z)$ by a scalar of the form $1 + cg$.

Remark: Operator product expansion in conformal field theory.

In a 2-dimensional conformal field theory, the OPE of composite operators has especially simple form. In this case, $V = \mathbb{C}$, and the space of functions on $V \setminus 0$ has a “bigrading”: the function $z^a \bar{z}^b$ has bidegree (a, b) ($a - b \in \mathbb{Z}$). The space of composite operators also has a bigrading: to any homogeneous operator \mathcal{O} one assigns two numbers – the holomorphic dimension $d(\mathcal{O})$ and antiholomorphic dimension $\bar{d}(\mathcal{O})$ ($d - \bar{d} \in \mathbb{Z}$). Therefore, if $\mathcal{O}(z)\mathcal{O}'(z') = \sum_k D_k(z-z')\mathcal{O}_k(z')$, then $D_k(z-z')$ has degree (d_k, \bar{d}_k) , where $d_k = d(\mathcal{O}) + d(\mathcal{O}') - d(\mathcal{O}_k)$, $\bar{d}_k = \bar{d}(\mathcal{O}) + \bar{d}(\mathcal{O}') - \bar{d}(\mathcal{O}_k)$. This implies that $D_k(z) = C_k z^{d_k} \bar{z}^{\bar{d}_k}$. Also, the action of the Virasoro algebra allows to reduce the problem of computing OPE of arbitrary operators to the problem of computing OPE for primary fields only, i.e. for fields which are highest weight vectors for the Virasoro algebra. In a rational conformal field theory, one has a chiral algebra of symmetries (for example, an affine Lie algebra) which is so big that there are only finitely many fields which are primary with respect to this algebra. This fact allows to treat a rational 2-dimensional conformal field theory in a purely algebraic setting, and reduce many of its problems to problems in algebraic geometry.

Appendix to Lecture 4: Remarks on renormalization and asymptotic freedom

Edward Witten⁴

December 1996

1. Ambiguity in operator products. In the last lecture we considered Lagrangians of the form

$$\mathcal{L} = \int d^4x \left(\frac{1}{2} (\nabla\phi)^2 + \frac{m^2}{2} \phi^2 + \frac{g}{4!} \phi^4 + \sum_i \varepsilon_i(x) \mathcal{O}_i(x) \right),$$

where \mathcal{O}_i are local functionals. We saw that defining correlation functions of such Lagrangians to the first order in $\varepsilon_i(x)$ (i.e. $\varepsilon_i \varepsilon_j = 0$) is equivalent to defining composite operators corresponding to \mathcal{O}_i . Now let us work to the second order in ε_i ($\varepsilon_i \varepsilon_j \varepsilon_k = 0$). This corresponds to considering products of composite operators. Before we considered products of composite operators at non-coinciding points x, y , and saw that such a product $\mathcal{O}_i(x) \mathcal{O}_j(y)$ is defined automatically once we have defined \mathcal{O}_i and \mathcal{O}_j . Now we will allow the points x and y to coincide. Then, as we know, the product is not automatically defined, and its definition requires additional renormalization. This means, we have to introduce counterterms in the Lagrangian, i.e. consider a new Lagrangian of the form

$$\mathcal{L}' = \mathcal{L} + \sum_{i,j,k} L_k \varepsilon_i(x) R_k \varepsilon_j(x) W_k(\Lambda, g) \mathcal{O}_k(x),$$

where L_k, R_k are differential operators (in x) with constant coefficients, and W_k are some functions which diverge as $\Lambda \rightarrow \infty$. The functions W_k are usually not uniquely determined and cannot be chosen canonically.

Let us consider an example. In one of the homework problems we computed the 1-loop correction to the 1-particle irreducible bosonic 2-point function Σ in QED. We discussed that this correction can be computed (in momentum space) as

$$\int e^{ikx} \langle J_\mu(x) J_\nu(0) \rangle dx,$$

where $J_\mu(x) = \bar{\psi} \gamma_\mu \psi$ is the operator of current, and the correlator in (1) is in the theory of a free fermion. Now, as follows from the above discussion, the product $J_\mu(x) J_\nu(x)$ is defined up to operators of lower order, and the expectation value $\langle J_\mu(x) J_\nu(0) \rangle$ is non-uniquely defined, which causes an ambiguity in the computation of (1). However, this non-uniqueness occurs only at $x = 0$. In fact, it is easy to show that the function $\langle J_\mu(x) J_\nu(0) \rangle$ is well defined up to adding a multiple of $\delta(x)$. Therefore, the ambiguity in (1) is a constant (i.e. is independent of k).

In QED, to preserve gauge invariance, it is necessary to choose this constant in such a way that the condition $k^\mu \Sigma_{\mu\nu}(k) = 0$ is satisfied. This gives a unique way to fix the constant.

2. Symmetry breaking. A symmetry that exists in a classical field theory may be lost in a particular renormalization scheme for the corresponding quantum theory. Of course it is possible that there exists a better scheme which preserves this symmetry, but it is also possible that there exists no such scheme, i.e. the symmetry is broken at the quantum level. For example, consider N free massless fermions ψ_1, \dots, ψ_N , with Lagrangian $\sum (\bar{\psi}_i, D\psi_i)$, $\psi_i \in S_+, \Psi_- \in S_-$. In this theory we have a $U(N)$ symmetry. Let us add interactions in such a way that part of this symmetry is preserved. For example,

⁴Notes by Pavel Etingof and David Kazhdan

add a gauge field A (i.e. regard the fermions together to make a section ψ of the vector bundle $S_+ \otimes E$, where E is an N -dimensional Hermitian vector bundle over the spacetime), which takes values in the Lie algebra of a subgroup $H \subset U(N)$, and consider the Lagrangian $(\bar{\psi}, D_A \psi)$, where D_A is the Dirac operator along the connection A . The classical symmetry of this theory is the centralizer $Z(H)$ of H . In quantum theory, however, this symmetry may fail for topological reasons. In other words, a topological anomaly may appear.

3. An oversimplified version of experimental confirmation of asymptotic freedom.

Consider a field theory with electromagnetic and strong interactions, which contains electrons (which interact only electromagnetically), and quarks (which interact both electromagnetically and strongly). The Lagrangian of such a theory can be written as follows. The fields are:

- (i) An $SU(3)$ -connection A_g (the field of strong interactions),
 - (ii) A $U(1)$ -connection A_e (the electromagnetic field),
 - (iii) Quarks q_i and an electron ε (which are fermions with values in \mathbb{C}^3 and \mathbb{C} respectively).
- The Lagrangian is

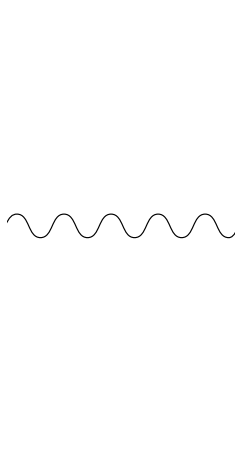
$$\int d^4x \left(\sum_j \bar{q}_j (iD + A_g + A_e - m_j) q_j + \bar{\varepsilon} (iD + A_e - m) \varepsilon + \frac{1}{e^2} F_{A_e}^2 + \frac{1}{g^2} F_{A_g}^2 \right).$$

Here e is the charge of the electron and g is the coupling of the strong interaction.

Now suppose that we scatter two electrons against each other with momenta p_1, p_2 , and measure the amplitude of the event that after scattering they will have momenta q_1, q_2 . As we know, this amplitude is defined by the 4-point function $\Gamma_4(p_1, p_2, q_1, q_2)$. Let us try to compute this function and thus predict the result of measurement.

First of all, we can use the fact that e^2 is small. This means, we can trust the perturbative expansion in powers of e .

To order e^2 , we can assume that the electrons, during scattering, exchange only one photon, which does not interact while it moves from one electron to the other. This corresponds to the following Feynman diagram:



Thus we have

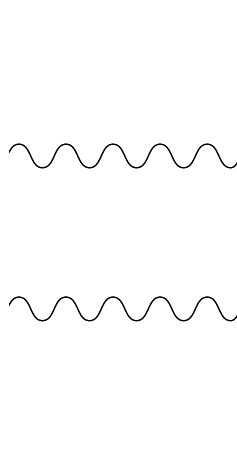
$$\Gamma_4(p_1, p_2, q_1, q_2) = e^2 (G_2((p_1 - q_1)^2) + G_2((p_1 - q_2)^2)),$$

where $G_2(p^2)$ is the free photon propagator. Here G_2 is regarded as an operator from $S \otimes S$ to $S \otimes S$, where S is the space of spinors.

Remark. In principle, we should include the diagrams where one of the electrons exchanges a photon with itself, but we will not consider them, regarding them as absorbed in the electron propagator.

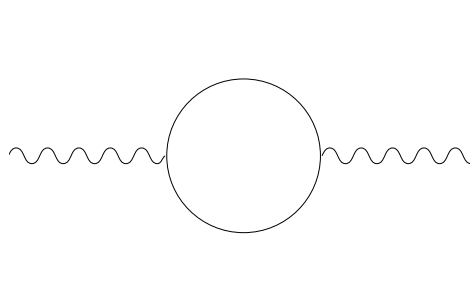
To order e^4 , we have two possibilities.

1) The electrons could exchange two non-interacting photons. The amplitude of the corresponding 1-loop diagram

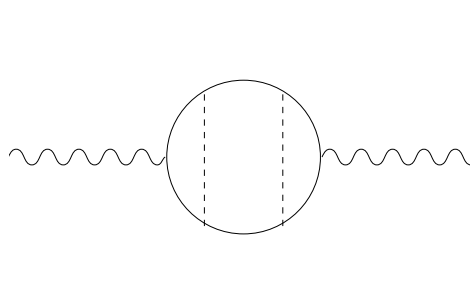


can be computed within the framework of QED.

2) The electrons could exchange only one photon, but on its way it could split in an electron and positron, or in a quark and an antiquark. The first splitting scenario

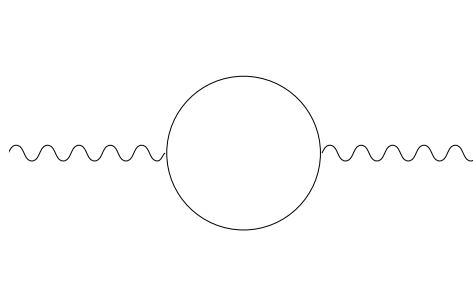


is harmless, since it gives only one 1-loop diagram with no strong interactions, and we can compute the amplitude of this diagram as in QED. However, the second scenario (with quarks) really gives us trouble. Indeed, the coupling constant g of the strong interaction is not small, so we cannot trust the perturbation expansion in g and thus have to take into account infinitely many Feynman diagrams with any number of loops:



(here dotted lines denote the field of strong interaction).

However, we know that the theory of strongly interacting quarks is asymptotically free (when the number of quarks is not too large). Thus, we should expect that the perturbative expansion in the effective coupling $g_{eff} \sim g(\ln p^2)^{-1/2}$ should be valid at high momenta. This would mean that at high momenta we can restrict to the 1-loop diagrams, which involve no strong interactions. Since we have one such diagram for each type of quark,



the total amplitude of these diagrams is $\sum e_i^2 \Sigma_i(p)$, where e_i are charges of quarks, and Σ_i are amplitudes of the corresponding diagrams for a particle with charge 1. Of course, the functions $\Sigma_i(p)$ depend on the (unmeasurable!) masses m_i of quarks, but at high momenta masses are irrelevant, and all functions Σ_i are approximately equal to each other. Thus, the amplitude is $\Sigma(p)(\sum e_i^2)$, where $\Sigma(p)$ is a universal function computed from QED (as in one of the homework exercises).

It follows from asymptotic freedom that the (relative) error of this computation is of order $1/\ln p^2$. If we compute the two-loop correction (still working to order e^4), we will get an additional term of the order $1/\ln p^2$, and the error will be of order $1/\ln^2 p^2$. More generally, if we take into account N -loop diagrams, the error will be of order $1/\ln^N p^2$. Thus, we get an asymptotic series, with very slowly decaying terms, but at very high p one can hope that it gives a reasonably good approximation to the 4-point scattering amplitude. This approximation (to the 0-th order) could in principle be checked experimentally, and can be regarded as a confirmation of asymptotic freedom.

Remark. In practice, asymptotic freedom was checked experimentally in a different way, but the ideology is similar to the one described above.

Correction to the text of lecture 3 (by Pavel Etingof)

Unfortunately, in Section 3.3 of Lecture 3 there is a wrong statement (noticed by D.Freed). Namely, the statement “ $\langle \dots \mathcal{O}(x) \mathcal{O}(x') \rangle \sim |x - x'|^{-[O] - [O']}$ ” is incorrect. For example, it fails if $O' = 1$ and \mathcal{O} is a nontrivial operator. This statement is true, however, if $O' = \mathcal{O}$, which is enough to make the point which was being made in the text.

Lecture 4: Scattering Theory

Edward Witten⁵

December 1996

Abstract

This lecture is an introduction to Scattering Theory in Quantum Mechanics. After introducing the asymptotic conditions and explaining their interpretation in terms of observable quantities, we introduce the Lippmann-Schwinger equation and the Born approximation. We then discuss Feynman diagrams and some differences between the non-relativistic and relativistic propagation of signals and particles.

1 Non-relativistic scattering theory

1.1 The asymptotic conditions.

We start by considering a particle of mass m moving in \mathbb{R}^3 in the time-dependent potential $V(\vec{x}, t)$ (whose support is not a priori assumed to be localized in the spatial directions). The evolution of the wave function $\psi(\vec{x}, t)$ is determined by the time-dependent Schrödinger equation

$$i\frac{\partial}{\partial t}\psi(\vec{x}, t) = H\psi(\vec{x}, t), \quad (1)$$

where the Hamiltonian is defined by $H = -\frac{1}{2m}\Delta_x + V$. We will assume for a while that $m = 1$.

We want to analyse the solutions of (1) with a certain behaviour in the far past and the far future. In the far past, we require that the solutions approach, as $t \rightarrow -\infty$, a solution of the Schrödinger equation corresponding to the *free* Hamiltonian $H_0 = -\frac{1}{2}\Delta$. Solutions of the free Schrödinger equation have the general form

$$\psi(\vec{x}, t) = \int \frac{d^3\vec{k}}{(2\pi)^3} f(\vec{k}) e^{i(\vec{k}\cdot\vec{x} - \frac{k^2}{2}t)}, \quad (2)$$

where $f(\vec{k})$ represents the distribution of momenta at time $t = 0$. More specifically, we want to consider those solutions of (2) corresponding approximately to an incoming particle with velocity \vec{v} , which means that $f(\vec{k})$ is localized near $\vec{k}_0 = \vec{v}$ (since $m=1$); a possible choice for f is

$$f(\vec{k}) = \exp\left(-\frac{1}{\alpha}(k - k_0)^2\right) \quad (3)$$

with small α . For large t , the integral in (2) is highly oscillatory so the main contribution comes from stationary phase; varying the exponent with respect to k gives $\vec{k} = \vec{x}/t$; if α is small then \vec{k} has to be close to k_0 , and the condition $\vec{x} = \vec{k}_0 t$ does indeed describe a free moving particle with velocity \vec{k}_0 .

In the far future, we can no longer expect the wave function ψ to be localised near a definite momentum because of the scattering effect of the potential, so we will instead look for solutions of (1) which behave asymptotically like a sum of a localised solution of (2) and an outgoing spherical wave.

In order to find solutions of (1) we will first solve an eigenvalue problem for the Hamiltonian. We assume that there are no bound states, i.e. that H has no discrete spectrum. For each \vec{k} we will search

⁵Notes by Radu Constantinescu

for a solution $\psi_{\vec{k}}(\vec{x})$ of the equation $H\psi_{\vec{k}}(\vec{x}) = \frac{k^2}{2}\psi_{\vec{k}}(\vec{x})$ with the following asymptotic behaviour at large distances:

$$\psi_{\vec{k}}(\vec{x}) \xrightarrow{x \rightarrow \infty} e^{i\vec{k} \cdot \vec{x}} + \frac{1}{|x|} e^{i|k||x|} F_{\vec{k}}(\Omega) + O\left(\frac{1}{|x|^2}\right). \quad (4)$$

The plus sign in the exponent of the second term is important, it will be needed in the construction of an *outgoing* wave solution of (1). The complex-valued function $F_{\vec{k}}$ is defined on the sphere of radius 1 and the notation Ω is a shorthand for a vector \hat{x} on the sphere. Alternatively we can regard $F_{\vec{k}}$ as a function $F(\vec{k}', \vec{k})$ defined for $|\vec{k}'| = |\vec{k}|$.

The functions $\psi_{\vec{k}}$ can be used to construct exact solutions of the time-dependent Schrödinger equation, namely

$$\psi(\vec{x}, t) = \int \frac{d^3\vec{k}}{(2\pi)^3} f(\vec{k}) \psi_{\vec{k}} e^{-i\frac{k^2}{2}t}. \quad (5)$$

Let us use (4) and (5) to see how does ψ look at infinity if f is assumed to be of the form (3). For $|t| \rightarrow \infty$ and $|x| \rightarrow \infty$ two terms contribute to the stationary phase; one of them is $\exp i(\vec{k} \cdot \vec{x} - k^2 t/2)$ and has already been analysed– the stationary phase condition is $\vec{x} = \vec{k}_0 t$. The other one has, for large $|\vec{x}|$, the phase $|\vec{k}| \cdot |\vec{x}| - k^2 t/2 + \arg F$; varying with respect to $|k|$ we find that the stationarity condition is

$$x = kt + \frac{\partial}{\partial k} \arg F_{\vec{k}} \quad (6)$$

Notice that if we ignored the term involving F we would get $x = kt$, which cannot be satisfied for negative t – in agreement with the fact that we only want to have an *outgoing* spherical wave.

The extra term in the right-hand side is a finite time advance or delay describing the fact that the moment when the scattered particles arrive at a detector is affected by the interaction with the target. The stationarity condition shows that the advance/delay is the radial derivative of the phase of F , called the *phase shift*.

1.2 Relation with experiments.

It is now time to relate the above constructions to quantities which are actually measurable in scattering experiments. The function $F(\Omega)$ is called the *scattering amplitude*. The reason for this is that the measure $|F(\Omega)|^2 d^2\Omega$ on the sphere $S_1 = \{|x| = 1\}$ can be interpreted as the number of scattered particles per unit solid angle if there is one incident particle per unit area ($d^2\Omega$ denotes the usual measure on the 2-sphere). Notice that only the absolute value of F is relevant here. The integral

$$\sigma = \int_{S_1} |F(\Omega)|^2 d^2\Omega \quad (7)$$

is called the *total cross-section* and represents the total number of scattered particles (per unit incoming particle in unit area). As such, σ can also be viewed as the *effective cross-sectional area* of the target: if the target is thought to remove a fraction of the incoming particles of momentum \vec{k}_0 , then σ is the total number of removed particles if there is one particle per unit area. The relationship between σ and F is usually written as

$$\frac{d\sigma}{d\Omega} = |F(\Omega)|^2. \quad (8)$$

1.3 The Lippmann-Schwinger equation.

We have so far used the eigenfunctions of the Hamiltonian to solve the time-dependent Schrödinger equation. The question now is how to solve the eigenvalue problem for H . One possibility is to start from the free Hamiltonian $H_0 = -\frac{1}{2}\Delta$ whose eigenfunctions are $\exp(i\vec{k} \cdot \vec{x})$ (corresponding to the eigenvalues $k^2/2$). We will search for eigenfunctions $\psi_{\vec{k}}$ of H of the form $\psi_{\vec{k}} = \exp(i\vec{k} \cdot \vec{x}) + \alpha_{\vec{k}}$. Since $H = H_0 + V$ we have

$$0 = \left(H - \frac{k^2}{2}\right)\psi_{\vec{k}} = V\psi_{\vec{k}} + \left(H_0 - \frac{k^2}{2}\right)\alpha_{\vec{k}} \quad (9)$$

which implies

$$\alpha_{\vec{k}} = -\frac{1}{H_0 - \frac{k^2}{2}}V\psi_{\vec{k}}, \quad (10)$$

where the operator $1/(H_0 - k^2/2)$ is supposed to be a right-inverse to $H_0 - k^2/2$.

Although the operator $H_0 - k^2/2$ is not invertible, we still can construct a right-inverse of it In momentum space, for instance, the operator $H_0 - k^2/2$ is roughly speaking a multiplication operator

$$\left(H_0 - \frac{k^2}{2}\right)e^{i\vec{q} \cdot \vec{x}} = \frac{1}{2}(q^2 - k^2)e^{i\vec{q} \cdot \vec{x}} \quad (11)$$

so we should have

$$\frac{1}{H_0 - \frac{k^2}{2}} = \frac{1}{\frac{1}{2}(q^2 - k^2)}. \quad (12)$$

Because of the pole, there are several ways of making sense of the above; one possibility, in the usual notation, is:

$$\frac{1}{H_0 - \frac{k^2}{2}} = \frac{1}{\frac{1}{2}(q^2 - k^2) - i\epsilon}. \quad (13)$$

The conversion to position space is realised through the following

Exercise. Show that the integral kernel (in position space) of the operator defined by (13) is

$$G(\vec{x}, \vec{y}) = \frac{1}{2\pi} \frac{e^{i|k||x-y|}}{|x-y|}. \quad (14)$$

By using the exercise, (10) leads to the *Lippmann-Schwinger equation*

$$\psi_{\vec{k}}(\vec{x}) = e^{i\vec{k} \cdot \vec{x}} - \frac{1}{2\pi} \int d^3\vec{y} \frac{e^{i|k||x-y|}}{|x-y|} V(\vec{y})\psi_{\vec{k}}(\vec{y}). \quad (15)$$

If the potential V has compact support then any solution of (15) has the boundary conditions required by the asymptotic condition built in. This can be seen by using the fact that for large $|x|$ and \vec{y} in a compact set we have $|x-y| = |x| - \hat{x} \cdot \vec{y} + O(1/|x|)$, where $\hat{x} = \vec{x}/|x|$. Indeed, if we use this approximation then (15) becomes

$$\psi_{\vec{k}}(\vec{x}) = e^{i\vec{k} \cdot \vec{x}} - \frac{1}{2\pi} \frac{e^{i|k||x|}}{|x|} \int d^3\vec{y} e^{-i|k|\hat{x} \cdot \vec{y}} V(\vec{y})\psi_{\vec{k}}(\vec{y}) + O\left(\frac{1}{x^2}\right), \quad (16)$$

which shows the existence of a scattered wave of the promised type.

Remark. The procedure used in (13) for going around the pole is chosen precisely in order to guarantee the existence of the outgoing spherical wave.

Notice that (16) gives an expression for the scattering amplitude: for any point $\Omega = \hat{x}$ on the sphere of radius 1 we have

$$F(\Omega) = -\frac{1}{2\pi} \int d^3\vec{y} e^{-i|\vec{k}|\hat{x}\cdot\vec{y}} V(\vec{y}) \psi_{\vec{k}}(\vec{y}). \quad (17)$$

Remark. We can reinterpret the function F as a function of k' and k such that $|k'| = |k|$, in agreement with the idea that after scattering the particle moves in the \hat{x} direction with the same speed as before.

Although the Lippmann-Schwinger equation cannot be solved exactly, it can be used to compute $\psi_{\vec{k}}$ in perturbation theory as an expansion in the powers of the potential (assumed to be sufficiently weak). This can be done by an iterative procedure: we first calculate $\psi_{\vec{k}}$ to order V by using the free eigenfunction $\exp(i\vec{k}\cdot\vec{y})$ in the right-hand side of (15) instead of $\psi_{\vec{k}}(\vec{y})$; once we know $\psi_{\vec{k}}$ to order V we plug it back into (15) to get the answer to order V^2 and so on.

1.4 The Born approximation.

We have seen above an expression of the scattering amplitude in terms of the eigenfunctions $\psi_{\vec{k}}$. The *first Born approximation* is the computation of the scattering amplitude with $\psi_{\vec{k}}$ replaced in (17) by the free plane wave $\exp(i\vec{k}\cdot\vec{y})$. We have seen that the scattered wave travels with the same speed as the incoming one so we can write, for $|\vec{k}'| = |\vec{k}|$,

$$F(\vec{k}', \vec{k}) = -\frac{1}{2\pi} \int d^3\vec{y} e^{-i\vec{k}'\cdot\vec{y}} V(\vec{y}) \psi_{\vec{k}}(\vec{y}). \quad (18)$$

In the first Born approximation this becomes

$$F(\vec{k}', \vec{k})_{Born} = -\frac{1}{2\pi} \int d^3\vec{y} e^{-i(\vec{k}'-\vec{k})\cdot\vec{y}} V(\vec{y}). \quad (19)$$

If we introduce the *momentum transfer* $\vec{q} = \vec{k}' - \vec{k}$ we see that $F(\vec{k}', \vec{k})_{Born}$ is the Fourier transform of the potential in the \vec{q} variable.

Remark. This fact is important since it shows that the less smooth V is, the less rapidly does $F(\vec{k}', \vec{k})$ decay as a function of $\vec{k}' - \vec{k}$ (for fixed \vec{k}). This led Rutherford to postulate the existence of the atomic nucleus (based on his scattering experiments). Later on, when similar experiments were performed with protons instead of atoms, the same reasoning suggested the composite structure of the proton (existence of quarks).

There are also higher Born approximations for the transition amplitudes: it is enough to use the successive approximations to $\psi_{\vec{k}}$ in the expression (17) of $F(\vec{k}', \vec{k})$. It turns out that in general the Born approximations lead to very precise computations.

Remark. Instead of considering a particle incident on a target we can regard our previous discussion as a description of the interaction of two particles. The center of mass decouples and the same results go through for the relative motion of the particles.

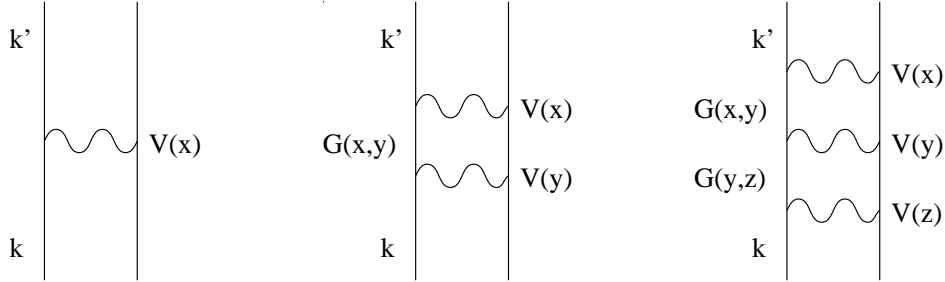
Remark. We can also generalize the preceding arguments for the case of n interacting particles. Let us note that this time we won't be able to simultaneously normalize the masses to 1 and that we need one spherical wave centered at each interaction point. The exact solutions of the Schrödinger equation $\psi_{\vec{k}_1, \dots, \vec{k}_n}$ will be sums of free and spherical waves away from the diagonals (i.e. whenever $\vec{k}_i \neq \vec{k}_j$ for $i \neq j$).

1.5 Feynman diagrams.

Starting from (17) and the Lippmann-Schwinger equation we can compute the transition amplitude as an infinite sum

$$\begin{aligned}
 F(\vec{k}', \vec{k}) &= -\frac{1}{2\pi} \int d^3\vec{x} e^{-i(\vec{k}' - \vec{k}) \cdot \vec{x}} V(\vec{x}) \\
 &+ \frac{1}{4\pi^2} \int d^3\vec{x} d^3\vec{y} e^{-i\vec{k}' \cdot \vec{x}} V(\vec{x}) G(\vec{x}, \vec{y}) V(\vec{y}) e^{i\vec{k} \cdot \vec{y}} \\
 &- \frac{1}{8\pi^3} \int d^3\vec{x} d^3\vec{y} d^3\vec{z} e^{-i\vec{k}' \cdot \vec{x}} V(\vec{x}) G(\vec{x}, \vec{y}) V(\vec{y}) G(\vec{y}, \vec{z}) V(\vec{z}) e^{i\vec{k} \cdot \vec{z}} \\
 &+ \dots
 \end{aligned}$$

Given our experience with Feynman diagrams, it is easy to see that the terms of the sum can be represented graphically (for instance, the graphs corresponding to the first three terms in the sum are shown below).



Each curly line represents an interaction due to the potential V . In-between interactions, the particles move freely (hence the free propagator $G(x,y)$). The incoming and outgoing particles have definite momentum (\vec{k}, \vec{k}' , respectively). The diagrams can be thought to encode either the relative motion of two particles or the evolution of one particle scattered by a fixed target. Notice that only very simple (ladder) diagrams appear, corresponding to that fact that there are no creation/annihilation phenomena in this non-relativistic description.

2 Relativistic versus non-relativistic scattering theory

We will now analyse some differences between the non-relativistic picture we have been considering so far and the relativistic treatment of scattering theory.

2.1 Propagation of particles

Going back to the time-dependent Schrödinger equation (1), we can try to solve it directly by imitating the method used in 1.3 to derive the Lippmann-Schwinger equation. We can use the plane wave solutions $\exp i(\vec{k} \cdot \vec{x} - \frac{k^2}{2}t)$ of the *free* Schrödinger equation to transform the Schrödinger equation for H into an integral equation: any solution of the perturbed Schrödinger equation satisfies the following analogue of Lippmann-Schwinger:

$$\psi = e^{i(\vec{k} \cdot \vec{x} - \frac{k^2}{2}t)} - \frac{1}{i\frac{\partial}{\partial t} - \frac{1}{2}\Delta} V\psi. \quad (20)$$

As in Subsection 1.3., the easiest way of making sense of the inverse of $i\frac{\partial}{\partial t} - \frac{1}{2}\Delta$ is in momentum space. Since

$$\left(i\frac{\partial}{\partial t} - \frac{1}{2}\Delta\right)e^{i(\vec{q}\cdot\vec{x}-Et)} = \left(E - \frac{q^2}{2}\right)e^{i(\vec{q}\cdot\vec{x}-Et)}, \quad (21)$$

an inverse (in momentum space) can be found by prescribing the way to go around the pole $E = q^2/2$. For instance, we could use

$$\frac{1}{i\frac{\partial}{\partial t} - \frac{1}{2}\Delta}e^{i(\vec{q}\cdot\vec{x}-Et)} = \frac{1}{E - \frac{q^2}{2} + i\epsilon}e^{i(\vec{q}\cdot\vec{x}-Et)}. \quad (22)$$

The integral kernel of the chosen inverse, in position space, is given by the inverse Fourier transform

$$G(\vec{x}, t; \vec{x}', t') = \int \frac{d^3\vec{q}}{(2\pi)^4} e^{i\vec{q}\cdot\vec{x}} \int_{-\infty}^{\infty} dE \frac{e^{-iE(t-t')}}{E - \frac{q^2}{2} + i\epsilon}. \quad (23)$$

The only pole of the E integral is at $E = \frac{q^2}{2} - i\epsilon$ and so, if $t - t' < 0$ we get $G = 0$ (because we can then avoid the pole by closing the integration contour in the upper half-plane). This result has an important implication: *particles can only travel forward in time*.

This is no longer true in a relativistic context: we have seen that the typical propagator of a particle of a mass m is

$$\frac{1}{q_0^2 - \vec{q}^2 + m^2 + i\epsilon}$$

in momentum space; in position space, the inverse Fourier transform gives (for $x = (\vec{x}, t)$)

$$G(x, x') = \int \frac{d^3\vec{q}}{(2\pi)^4} e^{i\vec{q}\cdot\vec{x}} \int_{-\infty}^{\infty} dE \frac{e^{-iE(t-t')}}{q_0^2 - \vec{q}^2 + m^2 + i\epsilon}. \quad (24)$$

No matter whether we close the integration contour in the upper or lower half-plane, we cannot avoid both poles, therefore it is no longer true that G vanishes if the time coordinates of the points x and x' satisfy $t < t'$. As a consequence, particles can make zig-zags in time, a phenomenon which is interpreted as the creation or annihilation of particle/antiparticle pairs (the particles traveling forward in time and the antiparticles backwards).

2.2 Propagation of signals

Non-relativistically, interactions are instantaneous. However, this is no longer true in the relativistic case.

Let us consider the example of the electromagnetic field; the interaction is transmitted by photons traveling at the speed of light (since the interaction is not instantaneous, we model it by some particles moving at finite speed).

The photon propagator in momentum space equals $1/(q_0^2 - \vec{q}^2 + i\epsilon)$, an expression whose non-relativistic limit (given by $q_0 \rightarrow 0$) is formally $-1/\vec{q}^2$. This has to be reinterpreted since non-relativistically there is no creation and annihilation of particles (so the only way we can think about the photon non-relativistically is to consider that it only exists at the time t_0 when the interaction occurs).

Notice that the (four-dimensional) inverse Fourier transform of $-1/\vec{q}^2$ equals

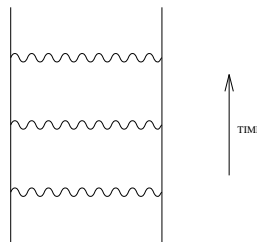
$$-\int \frac{d^4q}{(2\pi)^4} \frac{e^{iq\cdot y}}{\vec{q}^2} = -\delta(t) \frac{1}{|\vec{y}|}, \quad (25)$$

which is precisely a delta-function in time multiplied by the Coulomb potential. Therefore the non-relativistic limit corresponds to an *instantaneous* interaction (i.e. scattering in the Coulomb potential). Relativistically,

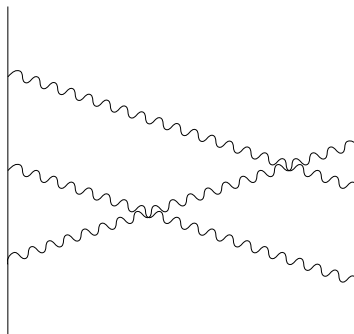
$$-\int \frac{d^4 q}{(2\pi)^4} \frac{e^{iq \cdot y}}{q_0^2 - \vec{q}^2 + i\epsilon} \quad (26)$$

is no longer supported at a fixed point in time, so the interaction is not instantaneous.

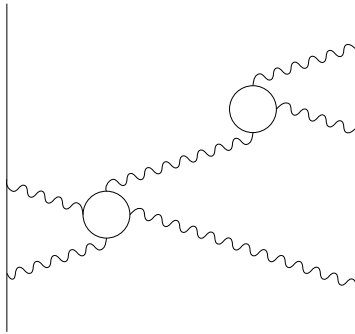
We can also illustrate the results on the propagation of particles and interactions in perturbation theory. It was shown that non-relativistically only ladder diagrams are encountered; intuitively, if time flows in the vertical direction, these diagrams represent particles moving forward with the horizontal curly lines being instantaneous interactions.



By contrast, in the relativistic case interactions travel at the speed of light and particle/antiparticle pairs can appear, so more complicated Feynman diagrams such as the ones below have to be considered.



This diagram illustrates the fact that photons traveling at the speed of light replace the non-relativistic instantaneous interaction. The curly line which represents the interaction in the non-relativistic case is relativistically a photon. In the second diagram we show the creation and annihilation of particle pairs. (Of course there are also diagrams in which both effects are present.) The diagram also illustrates another important fact: in a local theory, the presence of electron/positron pairs makes it impossible to count ‘the total number of particles in the universe’. The totality of electrons can be accounted for by a single electron zig-zagging in time or even a single closed loop.



The diagram also illustrates another important fact: in a local theory, the presence of electron/positron pairs makes it impossible to count ‘the total number of particles in the universe’. The totality of electrons can be accounted for by a single electron zig-zagging in time or even a single closed loop.

