# Renormalization

Rainer F. Hauser
rainer.hauser@gmail.com

March 29, 2020

**Abstract**

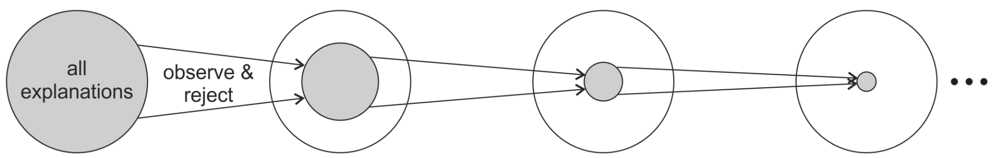
Because of divergent integrals in Quantum Field Theory renormalization became an important part of modern physics. It is also the theoretical apparatus to work with effective theories and get useful results without the need for the complete set of mathematical tools of today's physics. Tobias Osborne from the Leibniz University Hannover discussed it in the eighth video of his lectures which were available as "Advanced Quantum Field Theory" on YouTube at the time this transcript has been assembled and which may still may be available today. The lecture explains in detail what renormalization is, why it is relevant for doing physics, and how it is done in Quantum Field Theory.

## 1 Explaining the Universe

### 1.1 Progress in Physics

Physics starts in a first step with observations which lead to empirical data. In the beginning it is a list of unstructured data. In the second step one tries to explain the data because the list of data gets larger and more and more unmanageable. An explanation is a kind of data compression and is therefore easier to store. The goal of compression is not just reduction of the list of data but understanding as a third step. There are different ways in physics to get the impression that one understands the data. One possibility is a Hamiltonian $H$ (or Lagrangian $L$) plus a Hilbert space $\mathcal{H}$, another is a neural network which can reproduce the data, but still also a list of data somehow compressed may be a way of understanding. Physicists prefer the Hamiltonian plus the Hilbert space as a means of understanding, and the view of the universe in terms of quanta turned out to be the best way to explain all the collected data.

There is another part of physics that tries to make predictions in order to explain what might happen if one makes a new observation. One cannot confirm a hypothesis but only reject it when it turns out to contradict observations. At this point where one has proposed a hypothesis one has to go back to the first step and start again with collecting empirical data. This way physics contains a growing list of rejected models for the universe, and this process gives an increasingly smaller list of models $\{H, \mathcal{H}\}$ which explain everything. The set of all explanations is more and more reduced by rejecting those which contradict observations, and therefore with better observations the space of explanations becomes smaller and smaller. That is the process of doing physics.




As time passes and the space of explanations consistent with all observations gets smaller and smaller one starts to develop prejudices towards a certain explanation on the basis of simplicity. There is always one explanation that is consistent with all the observations ever performed. It is the list of all the observational data that has been collected, and this is a consistent and complete description, but it is not simple and does not give understanding. Simplicity is not easy to quantify, and it means different things to different people. Physicists look for a Hamiltonian and a Hilbert space, while a computer scientist probably finds a neural network simpler.

### 1.2 Simple Explanations

In physics simple explanations are models, and models are Hamiltonians that depend on some finite list of unknown parameters $\hat{H}_\Lambda(z_1, ..., z_n)$ with $z_k \in \mathbb{R}$ in some Hilbert space $\mathcal{H}_\Lambda$. (An infinite list of parameters would mean an infinite number of experiments to determine them all and does not make sense.) The Hamiltonian is here assumed to be a quantum Hamiltonian, and $\Lambda$ is the list of degrees of freedom to be explained. There is some evidence that $\Lambda$ in the universe is finite, but there is also some evidence that it is infinite. It is the beauty of physics that it is possible with a good model to explain an infinite number of degrees of freedom with a finite list of parameters.

As shown above physics start with the first step of making observations, and these observations give expectation values $\langle A_j \rangle = \alpha_j \pm \delta\alpha_j$ with $j = 1, ..., m$. In the second step one takes each model $\hat{H}_\Lambda(z_1, ..., z_n)$ not yet rejected and makes predictions for $\langle A_j \rangle$. If a particular value in $z_1, ..., z_n$ does not give the correct $\langle A_j \rangle$ then the model is rejected. In the end there is a map $f_j(z_1, ..., z_n, \Lambda) = \langle A_j \rangle$ with exact solutions (predictions). Usually this map is horribly many-to-one because there are multiple $z_k$ and $\Lambda$ that provide the same explanation for all the observations. It is a fact of life that there are multiple explanations that are consistent with all the observations made so far and there is at least one Hamiltonian that explains everything known today because the best model available is linearized gravity minimally coupled to the standard model with some cutoff. It is inconsistent with black holes, but one might argue the black holes have not really been observed. If the model is supposed to be as simple as possible (with a given notion of what this means) then one can select the correct model but what simple means may change over time.

The map $f_j(z_1, ..., z_n, \Lambda)$ with the predictions is not invertible because it is many-to-one. A notion of simplicity in physics around for a good hundred years is that a model $\hat{H}_\Lambda(z_1, ..., z_n)$ is simpler than another model if it requires fewer parameters and/or applies to more degrees of freedom. However, it may turn out in the future that Hamiltonians are not the right way to describe the universe and that this notion of simplicity is no longer applicable, but with today's concept of simplicity the ideal Hamiltonian depends on zero parameters and applies to all the degrees of freedom that are required to describe the universe. That would be the simplest model, and that is what string theory aims for.

### 1.3 Measuring the Parameters

These parameters $z_1, ..., z_n$ (the coupling constants in quantum field theory) may not be directly measurable, because it is not always the case that there is an experiment which measures $z_k$ directly. One makes a list of predictions and compares the predictions with the experimental observations. One therefore tries to invert the map $f_j(z_1, ..., z_n, \Lambda) = \langle A_j \rangle$ which is usually horrendously non-linear. The process of determining the parameters $z_k$ from the observations is therefore not operationally well-defined.

This is important if one wants to attach a notion of reality to a model. Thus, if one says that the coupling constant $\lambda$ is the scattering cross-section, for example, then one attaches an operational significance to $\lambda$ and is upset if it turns out to be infinity. It is perfectly legal that there are parameters for explaining the universe to be infinity. The inverse temperature of the universe, for example, is infinity if one believes that the universe is in a ground state. One can still make finite predictions. To summarize, all predictions $f_j = \langle A_j \rangle$ must be finite, and it should be possible to give finite predictions with infinite parameters $z_k$.

## 2 Renormalization in Quantum Field Theory

### 2.1 Cutoffs and Their Potential Problems

One encounters many infinities in quantum field theory. Some of them can be subtracted away because they are not operationally well-defined. An infinite ground state energy is an example since one can never observe the absolute ground state energy. One can shift it without any operational consequences (except in general relativity). Others such as scattering cross-sections are harder.

In the $\phi^4$ theory the Lagrangian giving many infinities is $\mathcal{L} = \frac{1}{2}(\partial_\mu \phi)^2 - \frac{1}{2}m\phi^2 - \frac{\lambda}{4!}\phi^4$ and it looks as if it has to be rejected. (What is discussed here for the $\phi^4$ theory can be generalized to any other quantum field theory.) When computing $f_j(m, \lambda)$ one encounters $f_j = \infty$ for fixed $m$ and $\lambda$ but these computations are done approximately and not exactly such that it is well possible that these approximations are the source of the problem. However, over the past fifty years these values have been derived in so many different ways that the general consensus is that these infinities do not come from the approximations.

Assuming fixed values for $m$ and $\lambda$ leads to infinities except for $m = 0$ and $\lambda = 0$, but the list $\Lambda$ of all the degrees of freedom one tries to explain has not yet been taken into account. These degrees of freedom are just all momentum modes so far. Since scattering is really taking place, the value $\lambda = 0$ can be excluded, but fixing $\lambda$ gives always infinity for the scattering amplitude. What is wrong here is the attempt to model for all degrees of freedom at once.

One should deal first with a smaller $\Lambda$ and impose a cutoff. The question is how to do this because the way to restrict the degrees of freedom is arbitrary. Another question is what the consequences of a cutoff are because the predictions $f_j(z_1, ..., z_n, \Lambda) = \langle A_j \rangle$ depend on $\Lambda$ and therefore on the cutoff.

## 2.2 Effective Theories

Because the parameters $z_k$ are not directly observable they may depend on the cutoff. If one can invert $f_j(z_1, ..., z_n, \Lambda)$ to get $z_k = z_k(\Lambda)$ such that the parameters can depend on the set of degrees of freedom and therefore on the cutoff, then one gets not one model but a list of models which depend on $\Lambda$ and are called effective theories. Thus, there is not the one true coupling constant but it is sufficient if one can find one coupling constant for the chosen cutoff which is consistent with all the observations because the believe that coupling constants have a direct operational interpretation has been given up. If one adds degrees of freedom one has to change these not directly measurable parameters in order to match all the predictions made so far.

Effective theories can work for all degrees of freedom one is ever likely to measure but do not work for all degrees of freedom theoretically possible. The order of limits here is important. One first imposes the cutoff and afterward works out the coupling constants as a consequence of this cutoff. This allows to make predictions up to this cutoff. If one wants to do predictions beyond this cutoff, one has to impose a new cutoff and recalculate the coupling constants. There is therefore a family of models.

The measurable quantities and the parameters are not the same. The electron, for example, has the mass measured as $9.1093837015 \cdot 10^{-31}$ kg but the quantity $m$ called mass in the Lagrangian is a coupling constant and not this mass. Only in the low-energy limit the coupling constant is the mass and, as stated above, coupling constants are not directly measurable but can only be inferred. At higher energies the measured mass and the coupling constant in the Lagrangian which is also called mass diverge.

Even with this relaxed condition that the parameters $z_k$ may depend on $\Lambda$ it is still far from obvious whether it can ever be satisfied or, in other words, whether one can by redefining a couple of parameters be able to explain a list of observational data as one changes $\Lambda$.

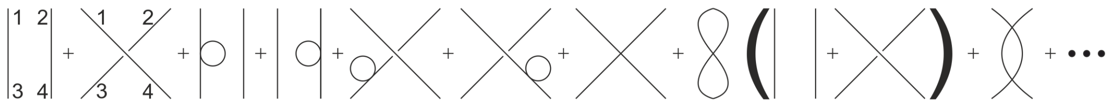
## 2.3 Renormalizable Theories

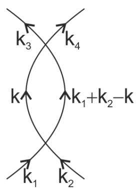
A theory which allows $f_j(z_1(\Lambda), ..., z_n(\Lambda), \Lambda) = \langle A_j \rangle$ for all $\Lambda$ and for a fixed number $n$ of parameters is called renormalizable or, in words, a theory which allows to absorb the cutoff $\Lambda$ into redefining the parameters $z_k$ but still matching a given list of data $\langle A_j \rangle$ is renormalizable. The fact that $n$ is fixed is important because the rest of the properties of renormalizability is trivial if the list of parameters needed to explain the data is allowed to grow. However, it is not clear whether this is always possible with a fixed $n$. Renormalizable theories are very remarkable objects in the space of explanations.

To do this for the $\phi^4$ theory one focuses on just one problem, the scattering amplitude. There is, of course, an infinite list of predictions a theory can make but it turns out that just by focusing on this problem one deals with all the other ones as well. This is an astonishing fact about quantum fields that by just focusing on the scattering amplitude (and possibly one or two other operationally well-defined quantities) one can already prove that a theory obeys the property of renormalizability.

## 2.4 Scattering Amplitude

With the S-matrix the scattering amplitude $\langle p_1\, p_2\, |S|\, p_3\, p_4 \rangle$ is the sum over all possible Feynman diagrams, and a Feynman diagram is just an integral. (The figure below shows the first few Feynman diagrams for the $\phi^4$ theory.) The result of this sum is infinity. By redefining the mass $m_0$ and the ground state energy $E_0$ some of the infinities can be eliminated.






The first infinity that cannot be eliminated this way for $|\Lambda| = \infty$ by redefining $\lambda$ similar to $m_0$ and $E_0$ is the one in the figure on the left side presented in momentum space. There is one undefined momentum $k$ in this diagram. The Feynman diagram in the figure corresponds to the integral

$$I = \frac{1}{2}(-i\lambda)^2 \int \frac{d^4k}{(2\pi)^4} \frac{i}{k^2 - m^2 + i\varepsilon} \frac{i}{(k_1 + k_2 - k)^2 - m^2 + i\varepsilon} \tag{2.1}$$

where the denominator of the integrand scales as $k^4$ and the integral is over $d^4k$ such that the divergence is logarithmic. The cutoff selected here is $|k| < k_c$ with $k_c \in \mathbb{R}$, and the integral $I$ becomes

$$I = \frac{1}{2}(-i\lambda)^2 i^2 \int_\Lambda \frac{d^4k}{(2\pi)^4} \frac{1}{k^2 - m^2 + i\varepsilon} \frac{1}{(k_1 + k_2 - k)^2 - m^2 + i\varepsilon} = 2ic \log\left(\frac{k_c^2}{(k_1 + k_2)^2}\right)$$

where $c$ is some constant one can work out but is not needed here.

The question is whether it is possible to adjust the parameters $z_1 = m$ and $z_2 = \lambda$ to eliminate the divergence of (2.1). The amplitude for scattering to order $O(\lambda^2)$ is

$$\mathcal{M}(k_c) = -i\lambda + ic\lambda^2 \left[\log\left(\frac{k_c^2}{(k_1 + k_2)^2}\right) + \log\left(\frac{k_c^2}{(k_1 - k_3)^2}\right) + \log\left(\frac{k_c^2}{(k_1 - k_4)^2}\right)\right] \tag{2.2}$$

and depends on $k_c$. If $\lambda$ is assumed to be one true value set at the creation of the universe, the only possible conclusion is that this quantum field theory has to be rejected. However, if one allows $z_2 = z_2(k_c) = \lambda(k_c)$ then there is a solution. To determine $\lambda(k_c)$ one assumes $\mathcal{M}(k_c, \lambda) = \mathcal{M}_{\text{exp}}$ with the value $\mathcal{M}_{\text{exp}}$ measured experimentally. If $\lambda$ follows $k_c$ according to the differential equation

$$k_c \frac{d\lambda}{dk_c} = 6c\lambda^2 + O(\lambda^3) \tag{2.3}$$

then it is possible to compensate the dependence on $k_c$ in (2.2) with the changing $\lambda$. This is legal because $z_2 = \lambda$ has no direct operational meaning. Thus, as long as $\lambda$ depends on the momentum cutoff $k_c$ in a fashion that obeys the differential equation (2.3) one can explain all measured scattering amplitudes once and for all.

This is just one example of how renormalization works, but the question remains what about other divergences coming from three-particle or four-particle scattering and all the other experiments one can perform. The surprising result is that one can eliminate all the infinities by only rescaling $m$ and $\lambda$. That is one of the most exciting results of quantum field theory.