Martin Zirnbauer

Renormalization — A Universal Tool of Modern Physics*

O. Introduction

The beginning of the 20th century saw several major breakthroughs in physics: Einstein formulated his theories of special and general relativity and Bohr, Heisenberg, Dirac and others unravelled the laws of quantum physics. These advances strongly shaped the path subsequently followed by physics, and they have also captured the fancy of some of the general public in spite of the fact — or perhaps because of the fact — that they describe phenomena far removed from our common experience and intuition. Einstein's theory of gravity is needed for the description of phenomena that occur over astronomically large length scales, i.e. over distances ranging from the radius of our solar system up to the radius of the entire universe. Quantum theory describes processes taking place at microscopically small length scales, in the mysterious world of atoms and subatomic particles.

Much has happened in physics during the six and a half decades that have passed since the heyday of quantum theory. Other major advances have been made that are comparable to those of Einstein and Bohr in their impact on physics: quantum field theory, spontaneous symmetry breaking, gauge theories, renormalization, critical phenomena, asymptotic freedom, fractals, spin glasses, chaos ... Each of these advances has brought a novel quality to physics by either opening up a new realm of reality or by changing our perspective of the consequences of existing theory. From the long list of major advances I have selected one topic, renormalization, for today's Fellow Colloquium. I made this choice because renormalization is an addition to modern physics with singular importance both for its internal consistency and for its predictive power.

For many a theoretical physicist, renormalization provides a calculational tool, a mathematical apparatus, enabling him to tackle a large variety of relevant and challenging problems that are too difficult to solve by any other known means. Needless to say, this practical and technical

* Colloquium held at the Wissenschaftskolleg, March 23, 1993.

aspect of renormalization is beyond the scope of what I can explain to an audience of the kind I am facing today. But in addition to being a calculational tool, renormalization in its modern formulation by Kenneth G. Wilson has become a general concept, a powerful framework in which to think — and sometimes speculate — about the properties of a large class of physical systems. Often it is not even necessary for the theoretical physicist to actually implement the cumbersome formal procedures of renormalization in order to make progress in his understanding. Rather, the mere knowledge that renormalization exists and operates in the way it does may be sufficient. More precisely, renormalization may guide the theoretical physicist in which aspects of a problem to focus on, it suggests relationships between seemingly unrelated physical systems, and it may lead to conjectures which can then be proved by detailed and sometimes painstaking work. It is these latter *conceptual* aspects that I consider to be of interest to a general audience, and I am confident that a fair idea of what they are can be conveyed to you in the sixty minutes I have.

Let me make it very clear right at the outset that I have contributed nothing at all to the development of renormalization. As a matter of fact, the modern view of the subject was already fully developed sixteen years ago when I began to study physics. The legitimacy of this talk derives from my fascination with the subject and from my having repeatedly been a *user* of renormalization in my past research. Moreover, I plan to continue to be a user of renormalization, mostly — as far as I can tell at present — in its application to the statistical theory of disordered systems, about which you have already heard a great deal in the talks of Hans Weidenmüller, Pier Mello and Axel Müller-Groeling.

My presentation will be divided into three parts:

1. The need for a cutoff
   (the catastrophe of infinities)
2. Renormalization at work
   — the general concept
   — an example
3. Filter action and some consequences
   — universality classes
   — change of perspective

The first part comes with the subtitle "the catastrophe of infinities". The notion of infinity I am referring to here is a mathematical abstraction, but it is an abstraction you are all familiar with. Pick a large number and add to it the number one to produce a number which is larger yet. Continue the process of addition and you will produce a sequence of numbers that grows beyond any bound. Increase beyond all limits is what implies the mathematical abstraction called infinity. Physicists — and, in fact, all natural scientists — use it constantly, for example when computing the derivative, or rate of change, of a smooth function. Nevertheless, the final outcome of any computation in physics, the value of any measurable quantity, such as the binding energy of the hydrogen atom, has to *be finite*, as opposed to infinite. In this latter sense, infinity has no place in physics. It therefore caused a severe crisis in physics when not just one but a whole multitude of infinities arose in the mid-thirties, as a result of physicists' attempts to combine Einstein's theory of special relativity with quantum theory to form what is called "quantum field theory". Because of its historical importance for the development of renormalization, I will try to give a hint of what the crisis was. By way of preparation, I will begin with another example that nicely illustrates the origin of the difficulty and has the virtue of being intelligible by common sense. It is an example that has become widely known through some popular textbooks on fractals.

1. The need for a cutoff

Consider the first drawing in Figure 1 which should remind you of the coastline of Britain, the largest of the British Isles. In fact, I made this drawing by opening the English section of a European road map with a scale of 1 : 10,000,000, selecting a sequence of points located on the British shoreline with a distance of about 100 miles from each other, and connecting the points by straight lines. Suppose now that, for some reason, we wish to know the total length of the British coastline. (We are the royal cartographers of a British monarch who has commissioned us to provide this information about the territorial extension of his lands.) For a first answer, we could take a ruler, measure the length of each of the linear pieces the drawing is composed of, add the lengths all up, and convert the result into the actual length by multiplying by 10,000,000. For the sake of the argument, let us assume that the outcome of this computation is 2,000 miles. The result so obtained is vulnerable to objection, however. After all, the British coastline is not straight over distances of 100 miles but is a highly structured curve, protruding into the sea and receding from it in irregular succession. This is indicated by the second drawing in Figure 1, which shows the segment between Fort Wrath and Ullapool magnified. So let us refine the estimate and base our computation on a more accurate representation of the British coastline, connecting by straight lines a sequence of points on the shore only 10 (rather than 100) miles apart. By the same procedure as before, we will then find that the result for the total coastal length has gone up, to about 4,000 miles or so. Continuing the process, we might turn to data (taken by a high-resolution satellite camera, for example) accurate down to one mile, a tenth of a mile, a hundredth of a mile, and so on, and we would obtain an ever increasing sequence of total coastal lenghts, 8,000 miles, 15,000 miles, 25,000 miles and so on.

It is now clear what I am driving at: the question posed (what is the total coastal length of Britain?) has no unique answer. In order to succinctly state the observations made so far, let the word "cutoff scale", or "cutoff' for short, mean the distance at which the linear approximation is applied. We can then say that the result for the total coastal length is *cutoff-dependent* and increases as the cutoff is lowered, i.e. as we measure it on a more and more refined scale. The rate of increase is a measure of the "roughness" of the coastline. If we assume coastlines to be rough on all length scales, even the very smallest, then the total coastal length grows beyond all bounds and becomes infinite as the cutoff is reduced to zero. You might argue, however, that such an assumption becomes meaningless at length scales below the size of a sand grain where the exact location of the boundary between land and sea is no longer well defined. Put differently, coastlines have a natural minimal cutoff, and from a pragmatic point of view we need not worry about infinity which arises here just because a mathematical abstraction is pushed to its limit.

The second example, which I will turn to now, gives more reason for worry. To begin, let me remind you that physics makes a distinction between matter and light, and that matter and light exert influences upon one another; we say that they *interact.* Matter produces light when forced to change its state of motion; a surface separating two different types of matter partially or totally reflects light; and light on its passage through matter may be absorbed by the matter and converted into heat. All these statements are easily verified by experimental observation. What is less easily done is to formulate a quantitative *theory* that explains the huge body of empirical knowledge about light and its interaction with matter now available. Twentieth century physics has produced such a theory; it is called QED (quantum electrodynamics).

In what follows I will first list some empirical facts basic to QED and then indicate how the combination of these facts led to a severe tangle with infinity and the need for introduction of a cutoff. First the basic facts.

(1) Matter consists of *elementary particles*, i.e. entities which are, for all that we know, indivisible. The property of an elementary particle that makes it interact with light is its *charge.* The most abundant negatively charged elementary particle, responsible for most of the interaction of matter with light, is the *electron.*

(2) Light, although it has fooled many eminent minds into believing it to

be a continuum with wave-like properties (my apologies to our Professor Häberle: Goethe, alas, was mistaken), consists of particles, too, a view already held by Newton and unequivocally confirmed by experiment at the beginning of our century. Light particles are called "photons" in physics.

(3) Photons and electrons follow the laws of quantum theory, which are probabilistic and prohibit our asking questions of too detailed a kind (Heisenberg's uncertainty principle). An example of a permissible question to ask is the following: given that an elementary particle was at point *A* at time t_1, what is the probability for the particle to be at point *B* at a later time t_2? To compute this probability we should consider all possible "paths" (two typical examples of which are shown in Figure 2) leading from *A* to *B* in time $t_2 - t_1$, and we should calculate a certain weighted sum over all of these paths.[1]

(4) Electrons, by their property of carrying charge, may both emit and absorb photons at any point in time and space.

One of the prominent consequences of principles (1) — (4) (with minimal further input of a technical kind) is a description of the process of scattering of photons by electrons. Consider Figure 3a, which shows a typical path of an electron first absorbing a photon at point *A* and then re-emitting a photon somewhat later at point *B*. This process of absorption and re-emission changes the energy and momentum of the photon. It is called "Compton scattering". Now, by principle (4), the electron may emit and reabsorb yet another photon *(p')* while on its way from *A* to *B*, see Figure 3b, and by principle (3), to calculate the probability for Compton scattering we must sum over all possible locations for absorption and re-emission to occur and over all possible paths taken by the electron and the photon *p'* in the intervening time. Here is the point where disaster strikes: the sum over intermediate paths turns out to be infinite! The mathematical reason for this is that what we are required to calculate by the laws of quantum theory is a sum over paths which are *rough on all length scales.* Thus there is a certain analogy to the British coastline, which became infinitely long when roughness was assumed to persist down to arbitrarily small length scales.

The difficulty is related to our notion of elementary particles as point-like objects: neither the photon nor the electron, to the best of our knowledge today, has a finite size that might serve as a natural minimal cutoff. Note that the appearance of difficulties upon assuming objects to be

1 What exactly the weight factors are in this sum I will not tell you since this would require making a mathematical excursion that would distract from the purpose of the talk.

point-like is a novel feature of quantum field theory not present, for example, in Newtonian physics. It is an excellent approximation, for the purpose of computing the orbit of the earth around the sun by Newton's laws, to idealize earth and sun as points. But even if this idealization were a bad approximation, it would still be *consistent*, in the sense that it would not give rise to difficulties of the kind that appear in QED.

Furthermore, the difficulty of infinities in QED is not easily discarded since it puts the predictive power of physics in jeopardy. It is true that for most practical purposes (let me mention here the theory of atomic structure, the theory of chemical bonding, and much of modern technology such as power generation, power transformation, telecommunication etc.) physicists do not need to resort to fully-fledged QED but can make do with simple approximate theories that, by virtue of their ignoring quantum theory (!), are free of the disease of infinities. Thus, the difficulty is more of a principal than of a practical kind, but in a science that claims to be basic to the other natural sciences, matters of principle are serious matters.

By 1948, about 15 years after the first appearance of infinities in QED, physicists had figured out a way of circumventing the problem. What they did precisely was to introduce a cutoff — as you would surely have guessed by now — and then to add so-called "counter terms" to eliminate those contributions that become infinite as the cutoff turns to zero. In this way, they turned QED into the quantitatively most successful theory of physics to date. The ad hoc procedure of introducing an arbitrary cutoff plus counter terms, however, left a bad feeling among many of the physics community since it gave the appearance of being a dirty trick that, in the words of CalTech physicist Richard Feynman, simply "brushed the problems under the rug". In fact, textbooks published as late as the mid-seventies declared QED and quantum field theory in general to be seriously flawed. All this has changed with the advent of the modern view of renormalization. I will come back to this change of perspective in the third section, after showing you "renormalization at work".

2. Renormalization at work

Having paid due reference to the historical origin, I will now turn to the main part of my talk, an attempt to explain what renormalization is and how it works. The first notion to be introduced is that of a "micro system", which involves the following:

(i) a number of "degrees of freedom" a_1, a,, ...;

(ii) a set of interactions $f_1, f_2, \ldots$ determining the influences exerted by the degrees of freedom upon one another;
(iii) a cutoff X guaranteeing that no infinities can possibly arise;
(iv) a "theory" *T, i.* e. a set of mathematical rules for computing from the degrees of freedom and their interactions the observable properties of the physical system they represent.

For example, the degrees of freedom might be atoms in a solid, the interactions the chemical forces acting between atoms, the cutoff the interatomic distance (for many purposes we may ignore the complicated internal structure of atoms), and the mathematical rules would come from the laws of solid-state theory. In my second example, the degrees of freedom are electrons and photons (jointly called "fields"), the interactions are defined by the charge and the mass of the electron (also called "vertices"), the cutoff is arbitrary, as long as it is small enough, and the theory is QED, i. e. an elaboration of the rules (3) and (4) listed in part 1.

Let me emphasize that our notion of micro system is very general. The degrees of freedom need not be subject to, say, the laws of quantum theory but may be describable in classical (or Newtonian) terms. Similarly, the interactions can be quite arbitrary in their form and strength, with just one exception: they are required to be *local.* What this means is best explained by way of our two examples. Atoms in a solid can establish chemical bonds only with neighboring atoms but not with atoms far away in space (Figure 4a). A photon can be gobbled up by an electron (by virtue of the electrons carrying charge *e)* only when their paths cross (Figure 4b).

A micro system has observable properties which, at least in principle if not in practice, can be computed from the theory, given the degrees of freedom and their interactions defined at a certain cutoff scale. Examples of observable properties are (1) the heat capacity of a solid and (2) the probability for Compton scattering in QED. Those properties of a micro system which are observable at length scales *much larger* than the cutoff scale are called the "macro properties". Physics aims at the often highly nontrivial goal of predicting these macro properties from a knowledge of the micro system only.

With all these preparations made, we now imagine in addition to the micro system $S_1 = (a_1, a_2, \ldots; f_1, f_2, \ldots; X; T)$ another micro system $S_2 = (A_1, A_2, \ldots; F_1, F_2, \ldots; A; T)$. The second one differs from the first one in that the cutoff is larger and the degrees of freedom are fewer. (The set $\{A_1, A_2, \ldots\}$ will typically be a subset of $\{a_1, a_2, \ldots\}$.) In Example 1, we might take for the cutoff *two* interatomic distances and select for the degrees of freedom the subset of white atoms in Figure 5a, which constitute only one fourth of the total number of atoms. In Example 2, we might require that the electron and photon paths to be summed over are straight over distances of $2^{x} 10'^{8}$ meters instead of 1×10^{-18} meters. Suppose now that the interactions F_1, F_2, ... have been chosen in such a way that the macro properties of the first and second micro system coincide. Then, since we are concerned with no more (and no less) than the prediction of macro properties, the second micro system is just as good for our purposes as is the first one. We say that the two micro systems are "macro-equivalent", and we call the second one a "renormalized" micro system.

You may wonder about the logic of all this, since our definition of renormalized micro system makes reference to the macro properties which, after all, it is our aim to compute! The crucial point is that physicists have established ways of constructing renormalized micro systems *without* going through the intermediate step of calculating the macro properties, provided that the interactions are local. Example 1: According to solid-state theory, the heat capacity of a solid composed of atoms vibrating from thermal excitation, is calculable from the so-called partition sum, of which we only need to know here that it is a certain statistical sum over the positions and velocities of all atoms in the solid. To construct the renormalized micro system S_2 what one does is to compute a partial sum, i. e. a sum over the positions and velocities of all the black atoms in Figure 5a, which are not part of S_2. (The sum for the white atoms is done later.) When this partial sum has been carried out, the black atoms have served their purpose — all they ever do is to vibrate and by doing the partial sum we have already taken this fully into account —, so they can be dropped and we arrive at the renormalized micro system (Figure 5b). The only remnant of the existence of the black atoms are forces acting between the white atoms, which we take for the renormalized interactions F_1, F_2, ... Similarly in Example 2, the renormalized interactions can be constructed by a clever way of organizing the sum over electron and photon paths.

In summary, renormalization is a function that assigns to every micro system S_i a renormalized micro system S_2 by increasing the cutoff and decreasing the number of degrees of freedom while adjusting the interactions so as to leave the macro properties unchanged. We denote this function by R, in formulae: $S, = (S_i)$. You will now ask why all this is useful and why, given the fact that the macro properties are already completely determined by the knowledge of S_1 alone, one should go through the process of constructing the auxiliary micro system S_2 at all. The usefulness of renormalization derives from the following features:

(1) depends only on the *change* of cutoff (i. e. on the ratio A / k) but not on the cutoff a, itself.

(2) can be *iterated* to produce a sequence of renormalized micro systems *S1* ->S2—>S3-*...—*Sn—*

(3) ' has the characteristics of *a filter*. (What this means will be explained in more detail later.)

These features make it possible to treat micro systems whose degrees of freedom conspire to interact in such a subtle manner as to make the computation of the macro properties extremely difficult if not impossible by any conventional technique. An example of such a "hard-to-treat" micro system, where renormalization is the only known method of solution, is given in the following. The example is *not* a standard textbook example but was chosen because it nicely illustrates all the important features while avoiding concepts that are unfamiliar to someone not trained in physics.

Let us consider two interfaces, and let us imagine that these interfaces exert a force upon one another (Figure 6 a). For concreteness, we take the first interface to be the boundary between a solid and a liquid, and the second one to be the boundary between the (same) liquid and a saturated vapor. (This choice, however, is not essential for what follows.) For such a physical system in equilibrium, there exist two distinct possibilities: the two interfaces may be bound to each other, or they may be unbound and dissociate. Suppose now that the forces acting between the two interfaces have been given to us and that the interfaces are in thermal equilibrium. It will then be our aim to predict which of the two possibilities is the one that is realized, the case of bound interfaces or the case of unbound interfaces.

The motivation for considering such a problem comes from the study of "wetting phenomena". Consider some vapor, say water in air, in contact with a solid surface. If the vapor is saturated and the surface cold, vapor particles may condense on the surface to form little fluid droplets (Figure 6b), a phenomenon vividly familiar to the Fellows living in the Villa Walter. We say that the liquid "partially wets" the surface. In addition, there exists the possibility for the liquid to make the surface completely wet by forming a continuous film as is shown in Figure 6c. (This occurs if the sum of the surface tensions of the solid-liquid and liquid-vapor interfaces does not exceed the surface tension of the solid-vapor interface.) In the latter case of complete wetting, one may ask whether the liquid film grows to some well-defined finite thickness or keeps growing thicker and thicker by continued adsorption of vapor particles. It is this question of stable versus unstable film thickness that leads, at a certain level of abstraction, to the problem of interacting interfaces formulated above.

I repeat the question posed: are the two interfaces bound to each other or do they dissociate? The answer is easy to give when both interfaces are *smooth*, and it is the following. The force between the two interfaces must be repulsive when they come very close to each other, in order to avoid touching and thus annihilation of one interface, a possibility which is ruled out by our assuming the vapor's readiness to condense on the solid.

If the force is attractive at some larger distance, then by continuity there exists some distance d_o where the force is zero and, in equilibrium, the two interfaces will be bound at this distance. If, on the other hand, the force is repulsive at all distances, then the interfaces will dissociate. These two alternatives are graphically depicted not in terms of forces but through the respective potentials in Figures 7a and 7b.[2]

The argument just given is qualitatively correct when both interfaces are *smooth*. However, we already know from the example of the British coastline, which can be viewed as an "interface" in two dimensions, that interfaces need not be smooth but may be rough under certain circumstances. This happens when gravity is weak or absent, and when the surface tension is sufficiently small, since then energy (favoring a smooth surface) is overcome by entropy (favoring an irregular surface which is statistically much more probable) to make the interface rough.

Let us now assume that the solid-liquid interface is smooth but the liquid-vapor interface is rough (Figure 8a). If the effects of roughness are small compared to the influence of the forces present, the situation is qualitatively the same as that analyzed before. Similarly, if the roughness dominates over the forces, the answer to the question posed is again simple: the interfaces dissociate. On the other hand, if the influences of roughness and attractive forces are comparable in size, the answer is far from obvious. Under these circumstances, we are neither able to compute the total force exerted by one interface on the other from simple considerations; nor can we base our prediction on considerations involving just a few atoms (i. e. a small part of both interfaces). The effects of roughness become operative only over lengths extending across many atomic units. Thus, what we are facing here is a physical system with very subtle collective behavior whose accurate prediction requires us to make calculations for a very large number of degrees of freedom. It is an example of a wide class of problems in physics known as "critical phenomena". In the absence of an exact solution, which is rarely available, the only known method of dealing with critical phenomena accurately is renormalization.

We shall model the problem of interacting interfaces by a micro system defined in the following way. The degrees of freedom a_1, $a2$, ... are the distances from the solid surface of the "interfacial points" marking the location of the liquid-vapor interface (Figure 8b). For the cutoff k we take the lateral distance between interfacial points and connect neighboring points by planar surfaces, which appear as straight-line segments in Figure 8b.

[2] Note that the "potential" of a force is a function with the property that its negative derivative equals the force.

There are forces acting between neighboring interfacial points and between interfacial points and the solid surface. The latter forces are given by a potential of the kind shown earlier in Figures 7a and 7b. Our theory is standard thermodynamics, which asserts that all equilibrium properties follow from the so-called canonical partition sum over all degrees of freedom (i. e. the interfacial distances in our case).

To construct the renormalization function we again organize the partition sum in steps. We first do the sum for all interfacial points colored black in Figure 8c while holding the white ones (whose sum is to be done later) fixed. In this way, the cutoff is doubled, the interactions are modified, and what emerges is a renormalized micro system (Figure 8d). We are particularly interested in the renormalization of the potential of the force acting between the interfacial points and the solid surface, since this holds the answer to the question whether the interfaces bind or dissociate. In Figure 9a, two different potentials are plotted (dashed and dash-dotted curves) together with the renormalized potentials produced by twenty applications of R (full curve). Two things are noteworthy here: (i) the renormalized potentials are purely repulsive telling us the nontrivial information that the interfaces dissociate in this case, and (ii) the renormalized potentials coincide. This is a general feature: all unbound micro systems of the kind considered renormalize to *one and the same* repulsive potential. Similarly, all bound micro systems renormalize to one of a one-parameter family of attractive potentials with increasing potential depth. (This is not shown in any figure here.) Finally, all micro systems that are critical, i. e. neither bound nor unbound, renormalize to the unique potential plotted as the full curve in Figure 9b. The dashed and dash-dotted curves in this figure are two examples of critical potentials.

3. Filter action and some consequences

We recapitulate: application of the renormalization function X' increases the cutoff of a micro system and, at the same time, adjusts the interactions in such a manner as to leave the macro properties, i. e. the quantities that are observable at scales much larger than the cutoff scale, unchanged. Notice that an increase in cutoff wipes out any detail of the forces acting below the cutoff scale. Because the number of degrees of freedom is reduced, the detail lost is irretrievable, and hence the process of renormalization cannot be reversed. This irreversibility is consistent with renormalization having the characteristics of a *filter:* some types of interaction are passed on to bigger cutoff scales with none or almost no change, others are completely held back or are substantially reduced. We saw a demonstration of this in the above example, where many applications of the renormalization function invariably gave rise to one of just three possible outcomes: a unique repulsive potential for every unbound system, a smooth attractive potential for every bound system, or the unique critical potential if the system is neither bound nor unbound. Thus, compared to the vast number of interactions that may occur in physical systems, it is only a very small number that actually passes the renormalization filter.

A micro system that survives many applications of the filter action with none or little change is called a "renormalizable field theory", or "renormalizable theory" for short. The set of micro systems which renormalize to the same renormalizable theory have the same macro properties. We say that they belong to the same *universality class.* It can be shown by rather straightforward techniques that the number of renormalizable theories and hence the number of universality classes is very small. This is the basic principle underlying universality in physics. Notice the tremendous gain in efficiency: instead of having to make a separate study for each of an infinite variety of possible micro systems, we need only enumerate the universality classes and study the properties of these. Often the universality class to which a specific micro system belongs, can be identified on the basis of symmetries (and other guiding principles) alone, without actually going through the formidable machinery of renormalization. This possibility of "short-cutting" renormalization makes the concept of universality classes especially useful in practice.

As a further illustration, let me add an example taken from my own field of research, the physics of chaotic and disordered systems. Axel Müller-Groeling has introduced you to the notion of a mesoscopic system as a physical system where quantum coherence plays a significant role. Recall that mesoscopic wires fabricated in the laboratory come with defects and various other irregularities, which we call "disorder". Axel showed you that the resistance of such a mesoscopic disordered wire exhibits random but reproducible fluctuations when some parameter like the strength of an applied magnetic field is varied. (Actually, Axel talked about rings, but his explanations apply equally well to straight wires.)

To get an analytical description of these fluctuations, Weidenmüller and his group in Heidelberg used a "coarse-grained" model, which divides the wire into compartments (see Figure l0a) of equal length, all of which, including the charge transport among them, are modelled by random matrices. When the results of the Heidelberg model were compared to those obtained from a more conventional model representing the disorder by randomly distributed impurities (Figure 10b), it was found that the results coincide! An explanation of this coincidence, which might surprise you in view of the rather striking differences between the two models, can be found by exploiting the filter action of renormalization. To do so, one formulates both models as so-called "supersymmetric field theories". (I am sure you do not wish to hear the details of this.) Upon increasing the cutoff, one in fact discovers that both models approach the same renormalizable theory and hence belong to the same universality class. Actually, a closer look reveals in this (as in the more general) case that a renormalizable theory need not be just a single point in parameter space, called a "fixed point" in technical language, but may consist of a whole family of systems characterized by a parameter called the "correlation length".

Instead of elaborating on the concepts of fixed point and correlation length, let me finish off by picking up on part 1 of my talk, where I anticipated that renormalization would change our perspective of QED. I am now in a position to explain in which way it has changed and why. Recall the need for introduction of a cutoff in QED to prevent observable quantities from being infinite. Such a cutoff is an element *external* to QED (and, in fact, to any quantum field theory), its magnitude is *arbitrary* and, worst of all, the observable quantities *depend* on it. Although there exists a way of cancelling the undesirable cutoff dependence, this requires treating the electron's mass and charge as free parameters not predictable by the theory. Not surprisingly, for a long time QED was widely perceived as being a theory not only incomplete but also seriously flawed. This has now changed because the filter action of renormalization suggests a plausible scenario that makes QED very acceptable to the elementary-particle physicist. The scenario is the following.

It is postulated that there exists some other, more fundamental theory of physics which *is* complete, i. e. free of arbitrary input, and whose observable predictions remain finite as the cutoff is lowered to zero. The more fundamental theory is supposed to describe physical phenomena correctly and completely at *all* length scales, including those below the Planck scale of 10^{-35} meters. This theory - I will call it the Grand Theory — is not known and, now comes the crucial point, *need not be known* if all we are interested in is a description of, say, atomic phenomena. At atomic scales, i.e. at distances of 10^{-12} meters or so, a cutoff of $10^{-}i^{5}$ meters offers a much finer resolution than is really necessary. We can therefore increase the cutoff to some subatomic scale by applying the process of renormalization. In this process, all of the wealth (or dearth) of information about the elementary constituents and the form and strength of their interactions in the Grand Theory is compressed into just two numbers, the mass and the charge of the electron, and what emerges is QED as we know it.

It must be admitted that such a scenario will have no direct consequences as long as the Grand Theory remains unknown. However, it does change our perspective by suggesting that QED is indeed incomplete, but incomplete for a good reason. It also changes the direction of future research by suggesting that QED is the best possible theory of atomic physics we could have hoped for and that any search for alternative theories is futile. Furthermore, since the remnants of the Grand Theory that are left over by the mutilizing filter action of renormalization are contained in no more than two numbers (the electrons mass and charge), QED is predictive. All we need to do is just to fix, once and for all, these two numbers by experiment. Every other observable quantity can then be predicted — within the limits of our calculational ability — from QED itself without any extraneous input.

I have argued that atomic physics is accurately described by QED (an incomplete theory) because phenomena observable at atomic scales are extremely insensitive to processes occurring at much smaller scales. The final point I would like to make is that this insensitivity implies a very pessimistic prognosis concerning the possibility of elementary-particle physicists uncovering the Grand Theory. To do so, they would have to figure out a way — and I don't see how they ever could as humans — of studying matter under the most extreme conditions, such as those that existed for a very tiny fraction of a second after the creation of the universe in the Big Bang.

The resulting uncertainty is reflected in a wide spectrum of differing opinions held by members of the community of elementary-particle physicists. On the one hand, there are those who expect the Grand Theory to be defined by a single mathematical formula of the greatest possible symmetry and simplicity; on the other hand, it has been suggested that the elementary constituents and their interactions be described by a truly vast random matrix of unimaginable complexity and irregularity. In my very personal and humble opinion, the question "is there order or chaos below the Planck scale?" will never be answered. By subjecting the forces of nature to the filter action of renormalization, God achieved two goals at once: He created atomic physics, governed only by the beautifully simple two-parameter theory called QED, but at the same time made sure that His knowledge of the most fundamental laws of physics will remain hidden from us forever.

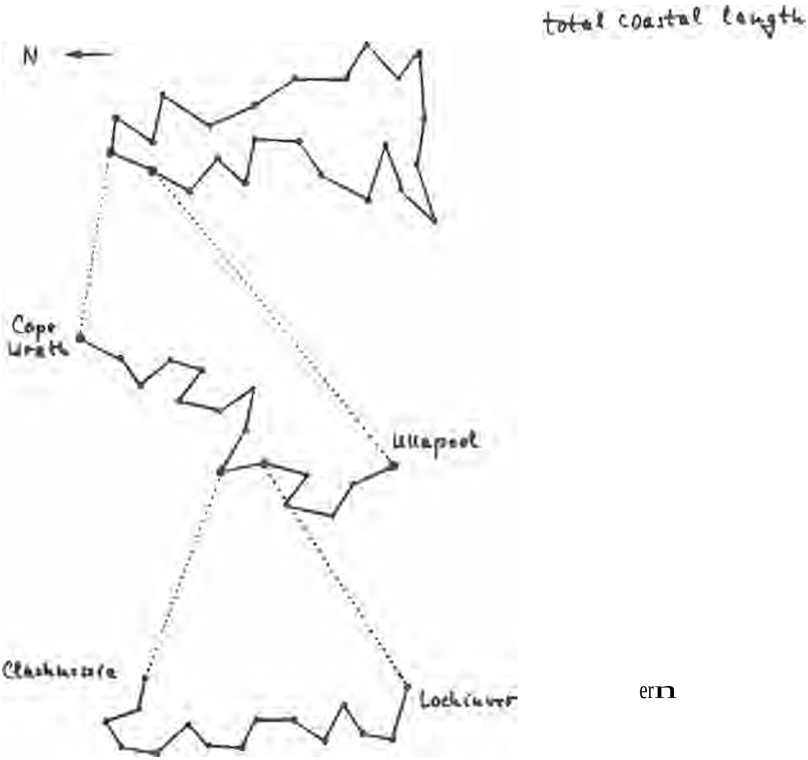


Fig. I: Roughness of the British coastline

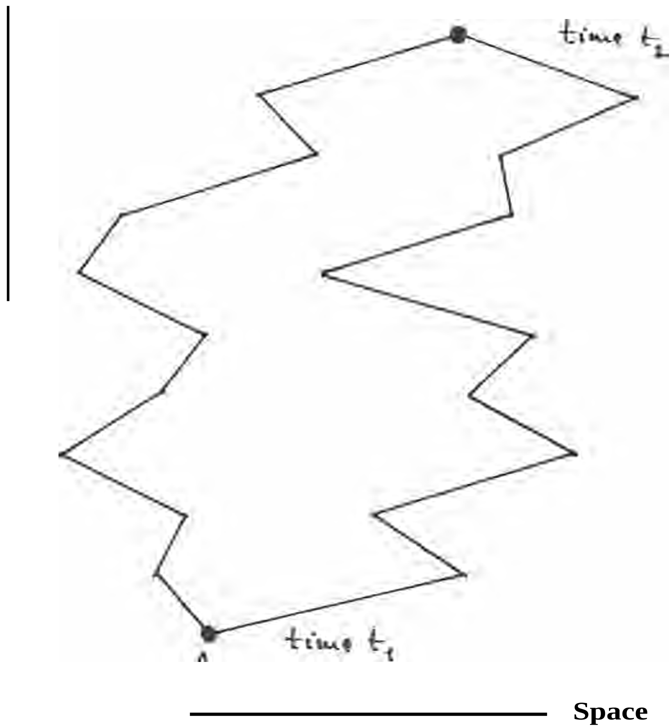


Fig. 2: Feynman's path sum in quantum theory

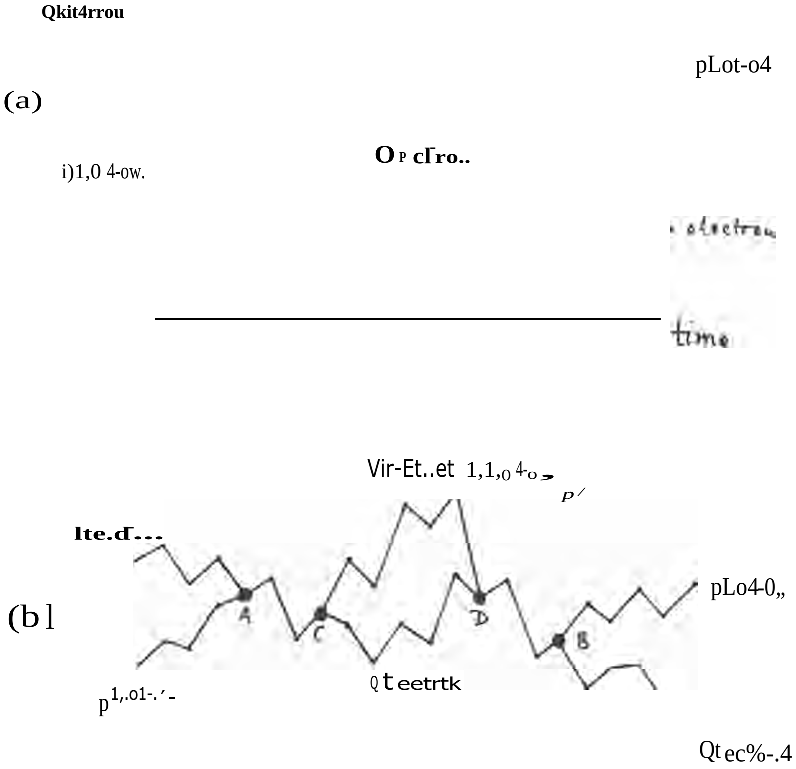
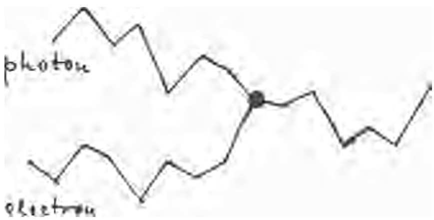


Fig. 3: Compton scattering

(a,

= et.ek..icat boket 6Q1- tti ht.:+kborev [ttok,r.

(b)

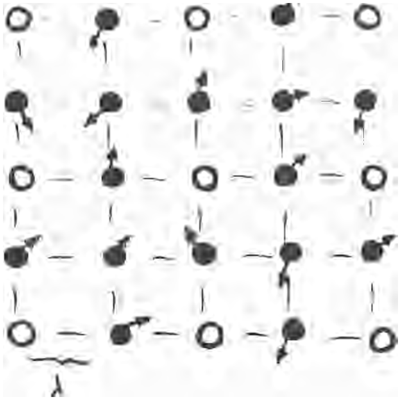


ɛt c.4⁻n>„

~ — Qttd,-bk - fi► ~otbk ~~~ez

Fig. 4: Local interactions

(a)



(6)

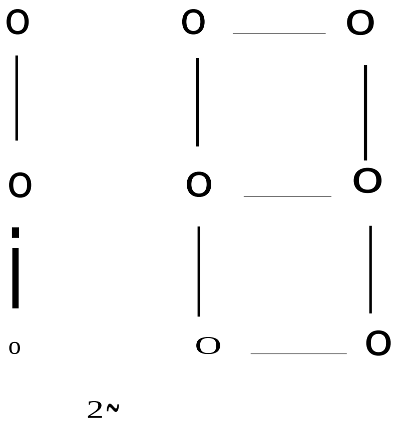


Fig. 5: Doing the partial sum

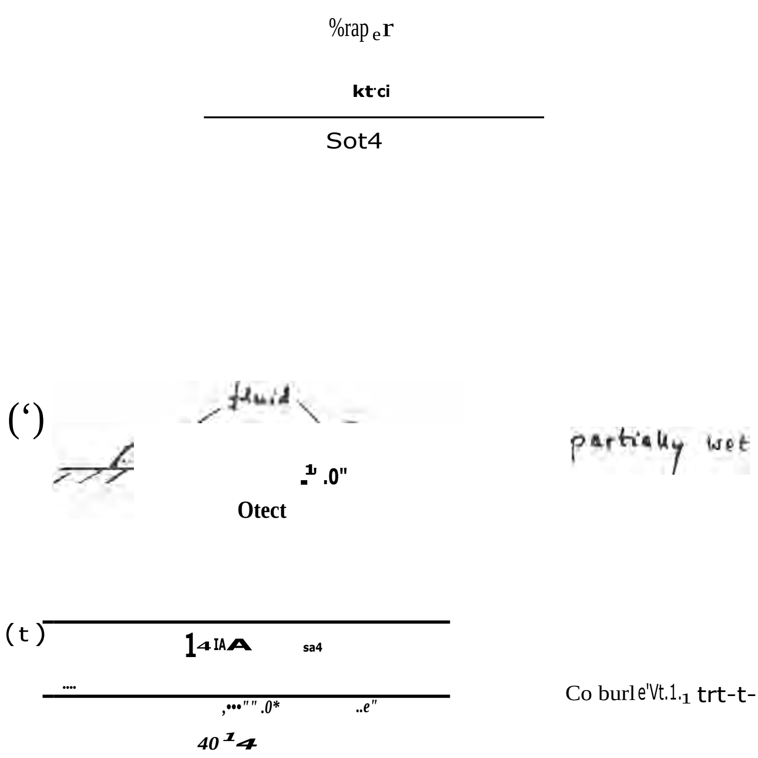


Fig. 6: Interacting interfaces

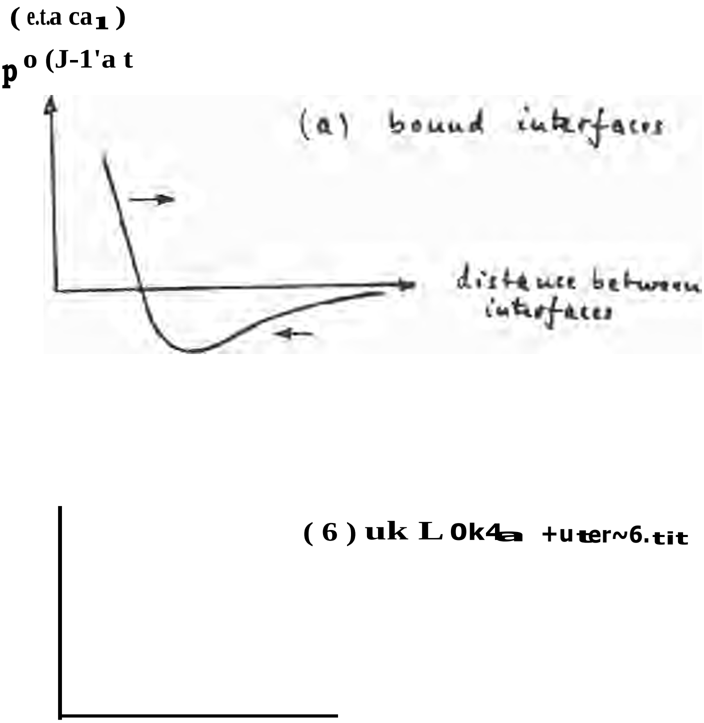
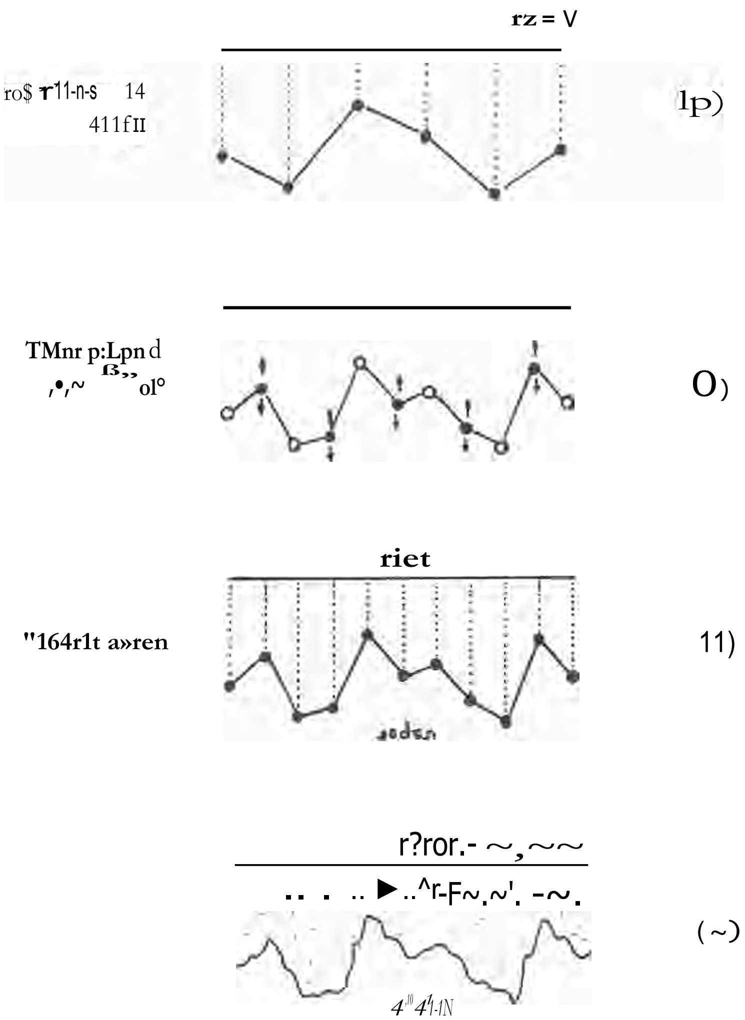


Fig. 7: Interface interaction potentials

aavpaiut gOnog :S ˙Fi,4



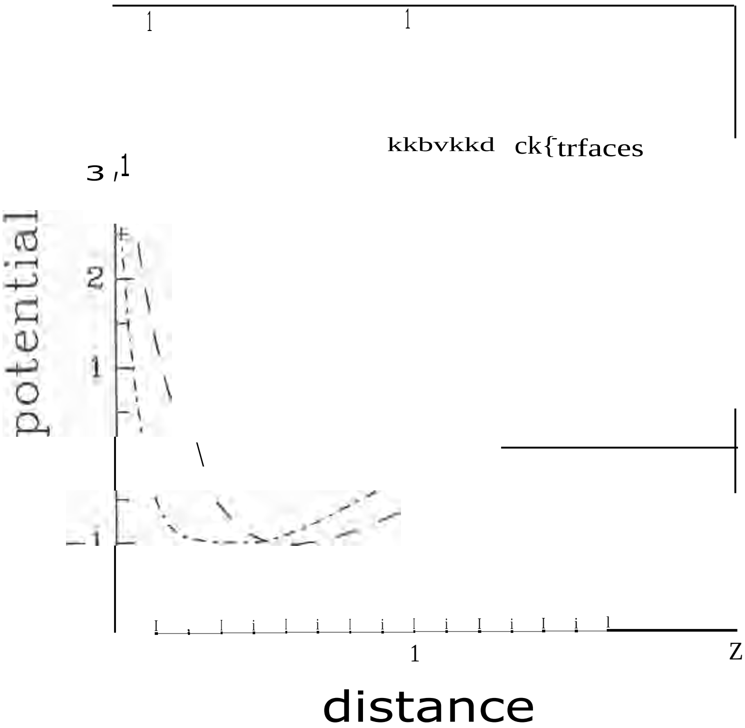


Fig. 9 a

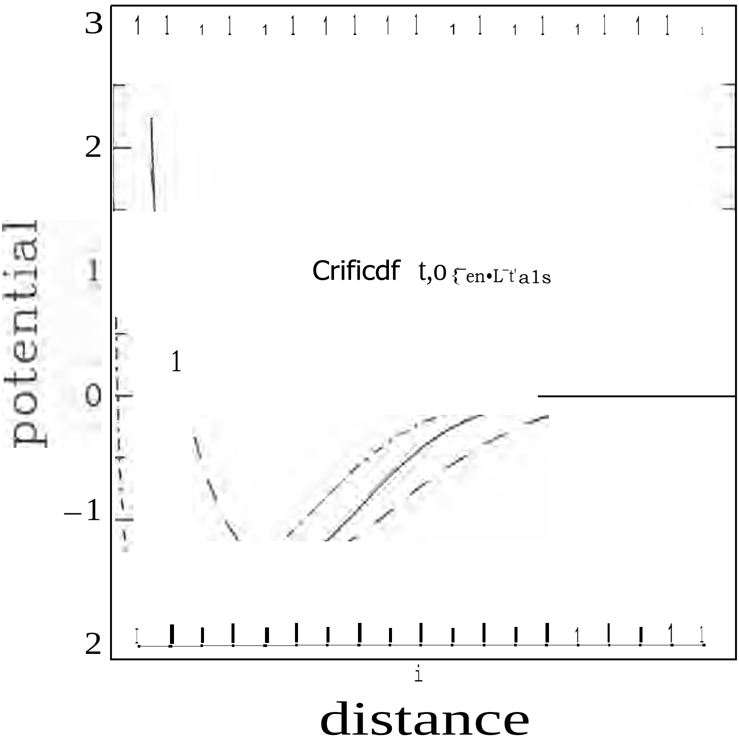


Fig. 9 b

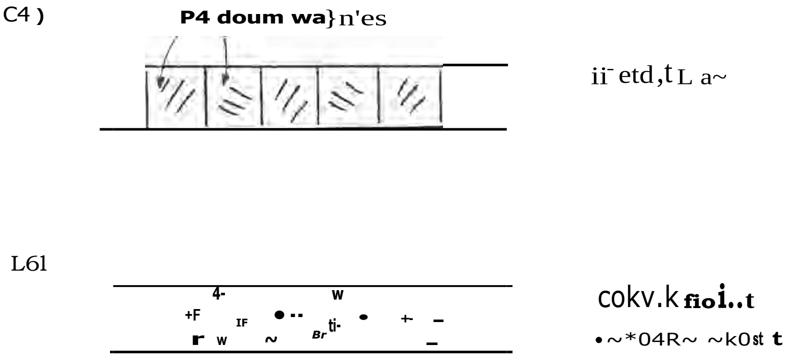


Fig. 10: Mesoscopic disordered wires