Investigating the Foundations of Physical Law
1 Introduction to foundational physics

What do we mean by Foundations of Physics?

Investigating the foundations of physics is a risky business. This is a subject without status, career structure, financial support, impact, prizes, journals (despite the existence of one naming itself *Foundations of Physics*), even protocols. For many people in physics, the subject simply doesn't or maybe even *shouldn't* exist. So why study it? The answer is simple: it's the cutting edge. It really is the frontier. That is why systems are not in place, why it has yet to find a place as an integral and essential part of physics. Because physics is a subject in which we are continually trying to push back to the 'origin' from a more complicated position, we can't lay down a generally-accepted basis until we reach the final discovery. When the final goal is to find the starting-point, you have a problem setting up a subject as conforming to the usual canons of scientific method, which is mostly structured in precisely the opposite direction.

Despite this apparent problem, we have to persist. Essentially, the most comprehensive model we have to date, the Standard Model of particle physics, has been brilliantly successful. But, as everyone recognises, it is only a *model*, not a theory. It gathers all the facts into a coherent structure, but without any explanation. And the Standard Model has been in place since 1973. We have had no major theoretical advance for forty years. What is the problem?

In my view, it is because we have confused the search for a *foundational* theory with the search for a *unified* theory. We have put our money on finding a combination of quantum mechanics with general relativity in a large overarching model like the Standard Model. We have imagined that some great superstructure, such as string or membrane theory, will somehow resolve everything, perhaps triggered by an experimental discovery that requires 'new physics'. But, even if we make new experimental discoveries, how will creating a complicated superstructure explain why we have such concepts as space and time to begin with? How will a 10-dimensional theory explain why we have dimensions at all? Even at a more basic level, how does the combination of space and time in both special and general relativity explain why space and time have fundamentally different properties? Combination theories are not asking the right questions. If we want to truly reconcile general relativity and quantum theory, we have to find the foundations on which they are structured.

Foundations of physics is not particle physics, or general relativity, or quantum mechanics, or any kind of extension or combination of them. It is a search for the explanations of such theories and the things that explain them. It is a search for the common origins of all physical theories, and their common origin with mathematics, in effect the explanation of the 'unreasonable effectiveness' of mathematics in physics and of physics in mathematics. If we make progress in this direction, then we will certainly be able to explain many things in classical physics, relativity, quantum mechanics and particle physics, and possibly areas of science of greater complexity. But this is like finding the technological consequences of blue skies research. We know that it will always happen, and we are happy to see it happen, but it is not the research's primary purpose.

How do we study it?

Foundations of physics is not only a separate discipline within physics. It also requires a completely separate way of thinking and methodology, one that is also intrinsically difficult and has yet to be included in any formal description of the scientific method. In principle, the method exists in that scientists have always used *induction* to infer a cause given certain consequences, but, in studying the foundations, this inductive approach needs to be taken to an extreme level. We have to find causes that are much more general than the ones we usually investigate, and this requires a much more imaginative and wide-ranging style of thinking than our training in the method of deduction from accepted starting-points, with perhaps a few carefully-limited inductive inferences, will allow. It means 'thinking outside the box' where the 'box' is physics as we have been taught to understand it. It *doesn't* mean trying to contradict this physics – a generally futile exercise – rather, it means trying to find something which may *look* different, and will certainly be simpler, but which automatically will give us the familiar structure when we make the connections. If we don't recognise this, then we have no hope of making progress.

Even if we do recognise it, there is no obvious way of going about implementing the process. Random speculation isn't likely to be successful if we don't have protocols. In effect, we have to have a theory of knowledge which is more fundamental than the knowledge we want to generate. We have to have a *philosophy*, a 'meta'-physics (in the sense of a more fundamental theory which explains how physics operates), perhaps even a 'metaphysics' in the usual sense. Now, it is quite common to hear physicists talk about 'philosophy' as the kind of thing you do when physics doesn't help you any longer – for example, how do we explain quantum mechanics? is there a God? etc. There is also a whole area of discussion about subjects such as time, the clock paradox, quantum correlation, which is described as 'philosophy of physics', and another area of discussion about how scientists actually operate called 'philosophy of science'. I am not talking about any of this when I say 'philosophy'. What I mean is strictly 'philosophy of knowledge', and strictly geared to working in Foundations of Physics. It is a highly technical process which aims to find those more fundamental principles of knowledge which will help us to choose fundamental principles for physics, and recognise them when we see them.

Now, no one would presume to state a set of general principles of knowledge and then hope that they would work for physics as we know it. The principles would have to emerge from some symbiotic process in which they were developed along with the physics in a continual feedback loop. If, however, the process eventually produced results that we recognised as being correct or significant, it would *then* be possible to begin to codify the principles for use in other cases. This is precisely the spirit in which I intend to propose to outline a methodology or 'philosophy' for investigating the Foundation of Physics, based on many years of working in this area.

The first thing is to try to get some idea of how we have actually already managed to develop a picture of the world which is very different from the one we normally perceive, and often contradictory to our expectations. To develop a capability of understanding nature, we have had to evolve as highly complex beings in an equally complex environment. We perceive the world at a level which is about 15 orders of magnitude bigger than the size of the proton and neutron, the main constituents of ordinary matter. Though even these are themselves structurally complex, the complexity they create in helping to produce the world as we perceive it is on a stupendous scale. We have no hope of ever reproducing it with any kind of exactitude. Complexity leads to *emergent* properties not seen at a less complex level, and these properties are the ones that we will naturally consider as 'normal', but hundreds of years of investigation have shown that we can be deceived. For centuries, one of the most certain ideas we had was that there was such a thing as solid matter and substance, but now we know that solidity and materiality are only emergent properties. Extended objects don't exist at the fundamental level; 'material' objects appear to be just a distribution of interacting points in an otherwise empty space.

The question is how did we manage to develop this view when our own experience of the world is so completely different? And the answer seems to be that we have been able to use the one real talent that we have – the one without which we couldn't have evolved as a successful species in the struggle for existence. Individual survival and procreation would have been impossible without *pattern recognition*, and the reason why it has been so useful to us is that nature seems to reuse the same patterns at different levels of complexity, a self-similarity and perhaps a fractal structure which is built into its very code. Because of this, we have been able to use patterns observed at one level to explain what occurs at another, as well as at the same time observing more sharply those aspects in which the pattern does not recur. Mathematics is a classic example of self-similarity, a kind of coding of the whole process. But the process has also been crucial in the development of the direct description of the real world in physics, the inherent self-similarity in physics being one of the reasons why we can use the mathematical coding.

Clearly, in looking at the Foundations, we need to see pattern as somehow a very significant component – and my own experience is that a highly-developed sense of pattern recognition is one of the most important components in 'thinking outside the box'. We also know from centuries of collective experience that mathematical structures must lie at the heart of physics, but that they must somehow reflect the transition from relative simplicity at the foundational level to more highly developed ones as more complex structures emerge from the foundations. It is no use producing a highly complex mathematical model as a *foundation*. It might lead to a few successful results in calculation, but, like Ptolemy's epicycles, it can't be the 'real' answer. On the other hand, whatever 'simpler' mathematical structures we use must be capable of leading by a natural progression to the more complex ones already in place. Somehow the mathematics and physics must be intimately connected – mathematics can't simply be a 'tool' used in physics. It has to be deeply built into the structures that physics needs for its foundations.

Simplicity is also key. The progression is always from simplicity at lower levels to complexity at higher levels, never the other way. Explaining complex systems using complex mathematics is an important art, and we have to be able to use it when the foundations are in place, but it doesn't help us to understand the foundations themselves. However, in looking for simplicity, we have a very significant clue which has emerged very strongly in the last few decades. One particularly significant type of pattern has become of immense significance, and it has to be an indication of how our mathematics can start from something seemingly simple and lead to something clearly complex. This is *symmetry*, which we see all about us in the laws of physics, the fundamental interactions and the fundamental particles. Finding symmetries can help us to decomplexify our explanations, and, if we can understand where symmetry comes from, lead to more profound understanding. We should also note that some symmetries are *broken*; that is, what is fundamentally symmetric appears, under certain conditions, to display some asymmetry. The reason for this cannot be arbitrary, and if we can discover it, along with the reason for symmetry, this will be a big step in our understanding of the foundations. One thing it can't be is *fundamental* because nature never acts in such an arbitrary way. It has to be, in some way, a sign of complexity or emergence.

Avoiding the arbitrary

In fact, avoiding the arbitrary is one of the cardinal principles of this way of thinking. Ultimately, you can't subscribe to physics only up to a point. It has to be the total and final explanation. There is no point at which you say 'this is where physics ends'. Of course, some people are on record as saying that there will never be a fundamental theory, and others as saying that physics might be different in different 'universes' – whatever they may be. I don't subscribe to any of this at all. There is no point in physics at all, in my view, if we don't believe that it has to be true without exception.

Believing in its unqualified truth has always been the source of its unique strength as an explanation of the universe. Whether we can ever find that truth is another question – though I am totally certain that we can.

Physics has been successful precisely because it does not compromise and this should suggest directions in which we should be looking for the foundations. If we decide a direction is the right one we have to be extreme in implementing it, utterly ruthless in pushing it to its limits. It is becoming increasingly clear, for example, that physics, in its foundations, must be totally abstract, exactly as quantum mechanics would suggest. We shouldn't, therefore, be spending time apologizing for quantum mechanics removing the idea of fixed material objects of a finite size 'interacting' with each other (whatever that may mean) from fixed positions in space. The idea that 'real' or 'tangible' 'objects' and purely abstract concepts like space and time can exist simultaneously in our physical picture is a logical impossibility – like having gods and humans in the same story. A foundational picture cannot be based on the simultaneous existence of things that are incommensurable. They have to be all of the same type, and all the evidence, including our experimental results on the point-like nature of fundamental particles, suggests that the 'true reality' is the abstraction, not the 'tangibility', which is only an emergent property at a higher level of complexity. Only complete abstraction makes logical sense, and only complete abstraction can bring about the desired link with mathematics. Quantum mechanics should, therefore, not be regarded merely as a 'calculating device', but as an exact indication of the way physics should go at the fundamental level. This is not only what physics is like, but what it *should be* like. My own investigations have indicated that a fuller understanding of the mathematical formalisms available to us for representing quantum mechanics provide a totally different understanding of what it is really about, and we will discuss this in the later lectures.

Again, if we decide that simplicity is to be preferred over complexity, we should be looking for something that is staggeringly simple, yet somehow capable of generating complexity. If we think our *basic* idea is a complicated construction, say a 10-dimensional space-time, then we have no way of knowing how this breaks up into the simpler component parts that must exist because we have no *fundamental* mechanism for doing this. We should certainly take notice of what the string theorists say about the symmetries required by nature, and we should expect to find them, even those expressed in 10 dimensions, but we should expect to find them by working out how such a complex idea emerges from simpler ones, in which the structures of the *components* reveal them as diverse in origin, the 'brokenness' of the larger symmetry coming from its inherent complexity, not by some arbitrarily-imposed concept of 'symmetry-breaking'. Broken symmetries are a sure signature of complexity, not of simplicity.

Once we start working purely with abstractions, *models* become meaningless. They are needed when we reach complexity because the complex structures are often difficult to understand without simplification and approximation, but no model lies at the foundations of physics, only pure abstraction. So, when we seek to reach this level, we automatically rule out model-dependent theories. At all times, we go for absolute minimalism, and abstraction is, in effect, minimalism in its final state. The famous rule known as Ockham's razor – we go for the idea which makes the fewest assumptions – is best exemplified by privileging ideas in which everything is there only for an abstract reason, with all arbitrary additions, made for our complexity-driven mode of comprehension (what I call the 'story-book' picture), removed.

Totality zero

There is one more extraordinary twist to the story. We have removed everything except physics. We have removed everything from physics except abstraction. Surely, we can now start from the abstractions. However, before we can do this, we must apply the principle of minimalism to the *abstract ideas themselves*. We ask the question: is nature at the foundational level characterized by them? And the only possible answer has to be no! Even the abstract concepts, simple as they are, are too arbitrary to be accepted as fundamental principles. The only possible statement that we can make is that nature cannot be characterized – it has no defining characteristics. Any characterization will necessarily be arbitrary. Here, now, we really need to think about philosophy. Is there a principle which can encapsulate the position we have arrived at, a metaphysical principle if you like? Fortunately, there is and it's been staring at us in the face in physics for quite a long time: *zero totality*. The universe and everything within it, including all the conceptualising that we can do about it, is absolutely nothing. This is the only position that we can take that makes logical sense. Complete zero is the only concept we can imagine that would not be arbitrary if unexplained.

The idea of zero totality is in no way contradictory to the way physics has been moving in the last fifty years. As long ago as the 1960s Richard Feynman was commenting on the seemingly remarkable fact that the negative gravitational energy of the Hubble universe effectively cancelled out the positive mass energy. He says: 'If now we compare the total gravitational energy $E_g = GM_{tot}^2/R$ to the total rest energy of the universe, $E_{rest} = M_{tot}c^2$, lo and behold, we get the amazing result that $GM_{tot}^2/R = M_{tot}c^2$, so that the total energy of the universe is zero. – It is exciting to think that it costs nothing to create a new particle, since we can create it at the center of the universe where it will have a negative gravitational energy equal to $M_{tot}c^2$. – Why this should be so is one of the great mysteries – and therefore one of the important questions of physics. After all, what would be the use of studying physics if the mysteries were not the most important things to investigate.' Cosmologists, similarly, have long spoken about everything starting from 'nothing', or zero space, time and matter. The well-known chemist and science writer Peter Atkins has said that 'the seemingly something is elegantly reorganized nothing, and … the net content of the universe is … nothing' (*Creation Revisited*, Harmondsworth, 1994, p. 23). Even classical physics is strongly based on the idea that the total force in the universe must always be zero – to every action there is an equal and opposite reaction.

How this would work out at a higher level of complexity would depend on how it was introduced at the foundational level, but at that level there is an obvious way in which it could happen; and this is something we have already discussed. If we are presented with a need to make everything total exactly zero then we have a ready-made explanation for why symmetry is so important at the deepest level. Symmetry (and, in particular, that form of symmetry known as *duality*) provides a clear way of having totality zero and yet allowing things to happen. This is routine for physicists, who, for example, will say that two objects will fly apart with zero total momentum, and yet each object has a nonzero momentum of its own. Now, if we want to introduce it as a *general* principle, we will have to imagine that the real zeroing happens at a subtle foundational level, and that apparent asymmetry or symmetry-breaking will come with complexity. This can be justified at a later date (and will be) when we look at how complex systems develop from simpler abstract notions, but, for the moment, we will simply accept that this is a principle which, like all the others, applies, at the foundational level, *without exception*.

What questions should we ask?

We now have in place our methodology. We require our foundational ideas to be simple, symmetrical, mathematically-based, minimal, totally abstract, and combining to a zero totality, and we want these conditions to be applied rigorously, ruthlessly, and in all cases without exception. We want to ensure that nothing enters into the picture which is in any way arbitrary, and that the final 'structure' (for want of a better word) is also *totally exclusive*, that is, that nothing in physics can exist outside of it. Anything breaking these conditions would not be truly foundational.

We have now established strict protocols for a foundational approach, all of which (even the 'metaphysical' principle of zero totality) are perfectly reasonable within the context of physics as we know it. This is only the first step – we haven't yet established what they will apply to – but it is a very important one. At this stage we want to exclude everything extraneous, to reduce the options to the only ones that are compatible with a foundational theory as we perceive it. The strictness is absolutely essential, as we are reaching the absolute limit of what it is conceivable to know.

Now, we need to establish what concepts fulfil these conditions. These will be the most primitive physical concepts imaginable. What we can do here is appeal to our original idea that the whole human development of science is only possible because we have the capacity to recognise repeated patterns. What we seem to find is that, as our investigation of nature descends from the large-scale complex structures to increasingly small-scale and less complex ones, certain concepts seem to be necessary at all levels (for example, space and time), while others (solidity, materiality) are ‘emergent’ and disappear at the less complex levels. This suggests that the former may appear at the most primitive level, while the latter will not. Using this kind of analysis we can put forward a provisional set of ‘primitive’ concepts (ones which cannot be broken down any further or discarded as emergent) which we can test against our protocols. The full account will be given in the third lecture after we have established some necessary mathematics.

However, it is clear that they must include space and time, but *not* space-time, which is emergent (otherwise we could not explain the differences between the components – another broken symmetry!), and certainly not curved space-time, which is even more emergent. Clearly, something must represent matter in its point-like state and something else energy or the connections between the points. There must also be some way in which, at the foundational level, these abstract concepts are delivered, all at once, as a ‘package’, explaining the growth of complexity in the way the packaging occurs.

Whatever we find won’t immediately look like physics as we now know it, any more than a cluster of cells looks like a human being, but, as we unravel the complexities, we will see that familiar physics will begin to appear, and with a greater clarity than it ever previously had. People have long had an expectation that some great complex equation will hold all of nature’s secrets, but nature doesn’t work like that at all. This primitive ‘structure’ will certainly be mathematical, but it won’t be an equation. Its secrets won’t be immediately obvious. They will have to be teased out one by one, as the complexity develops. This is physics at a primitive, embryonic level, an *Ur-Theorie* if you like. It requires inductive thinking, a kind of X-ray vision which penetrates through the layer upon layer of complexity to the primitive core – deductive mathematical techniques won’t help us. But the severity of the criteria we have established and the exact mathematical basis will ensure that it is rigorous. And it will begin to answer those questions, like what are space, time and matter?, why is space as we observe it 3-dimensional? and why does time never run backwards?, that we may have thought could never be answered.

Peter Rowlands
Physics Department, University of Liverpool

8 March 2013

Acknowledgements

The work in these lectures forms part of a long-term research project, whose most significant publication is a book entitled *Zero to Infinity: The Foundations of Physics*, published by World Scientific, 2007, although there have been a considerable number of earlier and subsequent papers and books related to the subject. Before 1997 the project was a solitary enterprise and much of it still is. However, since then, there have also been a number of collaborations, of which the most important (in chronological order) are those with John Cullerne, Bernard Diaz, Peter Marcer, Brian Koberlein and Vanessa Hill. Where work has emerged from one of these collaborations, this will generally be stated, but the details can be found in *Zero to Infinity*. I would also like to mention the continuing support from a number of members of Liverpool's Physics Department, and in particular from Mike Houlden and John Dainton.