The CONCEPT of the "FIELD" in PHYSICS

The idea of a 'field' in Nature has its origin in the much older ideas about the "aether", in ancient European and Middle Eastern cultures. The aether first appears in ancient mythology as an all-pervasive 'fluid' which was typically endowed with remarkable properties. In the philosophy of Aristotle the aether is the quintessence, or 'fifth element', existing only in the celestial regions of the universe; for more on this see the notes on Aristotle (as well as antecedent ideas amongst the pre-Socratics). Throughout the middle ages in Europe such ideas animated searches for traces of the quintessence on earth - it was widely believed that the possessor of such traces would be able use them to do things otherwise impossible, such as transmute substances into Gold, etc.

The first serious employment of the idea of the aether in physics came with Huyghens and Newton. Their reasons for doing so were rather specific to their own theories, and had little to do with historical ideas of the aether. In Newton's case, he needed the luminiferous aether to explain refraction and diffraction in his particle theory of light - for him the aether was an invisible medium quite distinct from both light and matter. Huyhgens, on the other hand, treated light as a wave-like oscillation of the aether; thus for him Nature was composed only of matter and the aether, and light was not a separate entity from the aether. Newton was also tempted to use the aether as a medium for the transmission of gravitational forces. Common to all these ideas of the aether, however, was the idea of a continuous medium which permeated all of space.

The modern physical idea of the "*Field*" began in the work of Faraday and Maxwell, who finally were able to elucidate the real physical nature of electrical and magnetic phenomena in terms of a single entity, the Electromagnetic (EM) field. Up to this point one could have argued that the EM field was just another kind of aether, albeit now with very precisely defined properties. However the advent of Special Relativity theory in 1905 made it clear that in the old sense of a medium in space, the aether did not exist - the existence of a 'frame of reference' defined by the aethereal fluid contradicted the principle of relativity (see notes on Relativity theory). The generalization of this theory in 1915 by Einstein, to include gravitation and matter, made the situation more confusing, for it seemed as though one could think, in a certain sense, of spacetime itself as a kind of field - this has led some, including even Einstein, to refer to some of the properties of spacetime as being similar to those of an aether. In modern physics, everything is considered to be a 'quantum field', including spacetime. However this idea is certainly provisional, since there is a deep contradiction between quantum mechanics and General Relativity, our two most basic theories, and this contradiction has yet to be resolved.

In what follows I first give a simplified description of what we mean by a field in the abstract sense (ie., without referring to any field in particular). I then briefly describe how the term is used in modern physics, and how this usage has developed in the last 2 centuries.

The MATHEMATICAL IDEA of the FIELD

It is important to understand what we mean by the idea of a field, quite separate from any specific example. Quite generally, a field is defined as some quantity which can vary continuously in some domain (usually in the domain of space and time). To see how this works, let's consider some examples.

(1) SCALAR FIELDS: A scalar field is one in which the quantity of interest, that is varying around the domain (eg., spacetime) is a *number* (the fancy mathematical term for a 'number' being a 'scalar'). One can think of many cases that everyone is familiar with. For example:

Altitude Fields: Suppose we consider some topography (eg., the topography of a region of the Gulf Islands, shown in Fig. 1(a)). Now we are all accustomed to a topographic map, showing the contours of equal altitude - this is what we see in the map. Imagine how one might go about producing such a map. We would measure the height of a very large number of different points in the region or *domain* shown; each point would be labeled by its exact position (eg., longitude and latitude) and given a height. This would not show the height as it varied continuously with position - for that we would need an infinite number of measurements at each and every point of the domain. It would be more like a chart of depth soundings, which show the depth at a finite number of different points on the sea. However if we had mapped out the altitude at a sufficiently large number of points on the map, we could *interpolate* between them to produce a continuous map of the altitude as it varies around the domain. We could then connect every single point having the same altitude to give a contour line. Obviously this map is only as good as the measurements. If, say, we measured the altitude on the ground at every point of a grid array, with neighbouring points on the grid separated by 1 metre, we would have a pretty good map, but it would miss features that were much smaller than a metre in cross-section (eg., a flagpole).

In any case, we see that the topographical map is a way of representing the altitude A as it varies continuously with position $\mathbf{r}$ on the map. We say that the altitude is a 'function' $A(\mathbf{r})$ of the position - if we know what $\mathbf{r}$ is, then we can say what the altitude $A(\mathbf{r})$ is at the point $\mathbf{r}$. This we can say that the altitude is an 'altitude field' defined in the domain covered by the map.

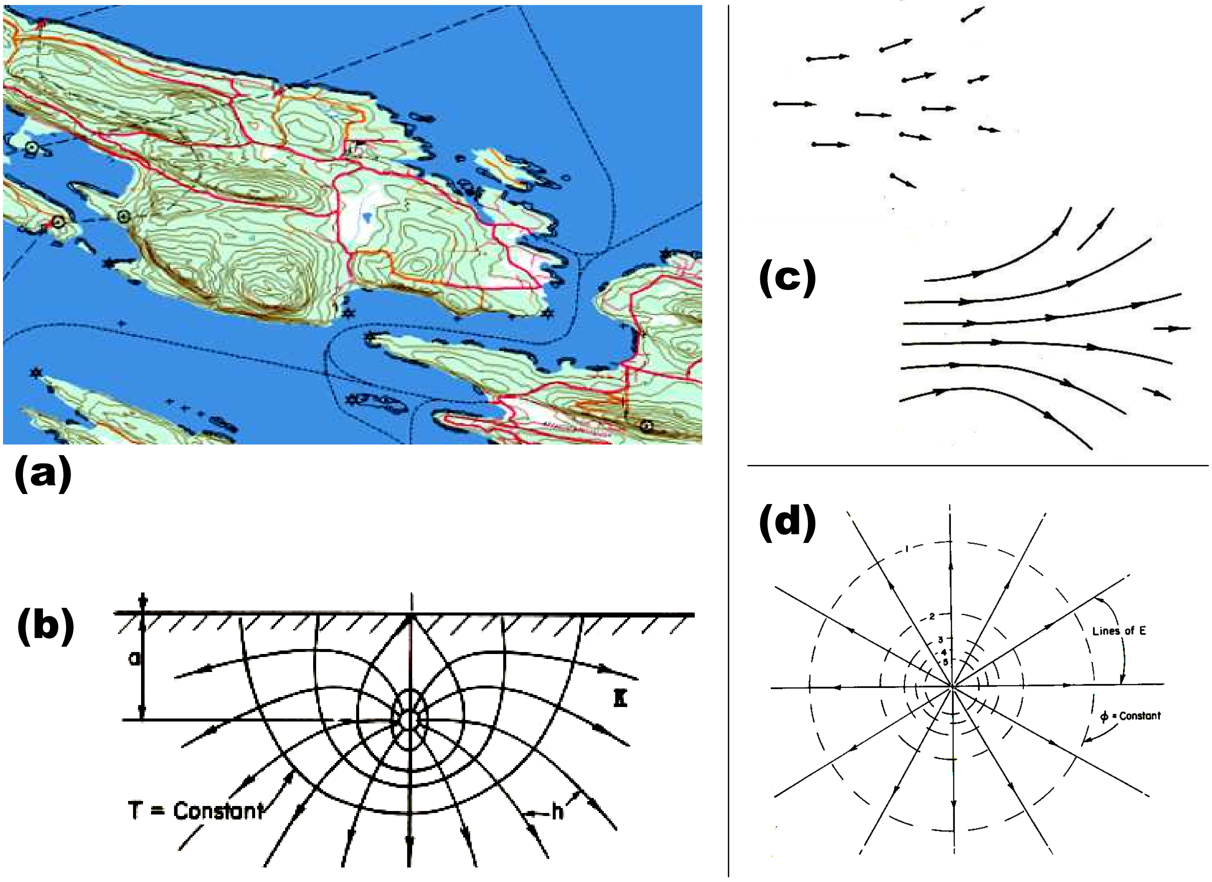

FIG. 1: The mathematical idea of a field. In (a) we see a contour map of the altitude on a section of the Gulf islands - contours show equal heights above sea level. In (b) we see a temperature contour map - for details see text. In (c) we see a vector field - the top left picture shows a map of the field at a few locations, and the lower right picture joins these into continuous streamlines. In (d) we see the electric field lines around a positive electric charge.

Temperature fields: Suppose we consider a room, and ask how the temperature varies around the room. There might be a stove in the room, so that the temperature is high around the stove and above it (because of the hot air rising above the stove), but much lower near the floor. At every point in the room we suppose that the temperature has some value, which could in principle be measured by a thermometer. Thus we see that the temperature is just some number (ie., a *scalar*) at every point in the room - and the space in the room in which the temperature varies is just the space in the room. Thus we can say that the temperature T takes some value at every point in the room. If we label different points in the room by their position $\mathbf{r}$, then we can say that the temperature T is a 'function' of $\mathbf{r}$, which we call $T(\mathbf{r})$; this just means that for every different possible value of $\mathbf{r}$ (ie., for every different position in the room), the temperature takes some value $T(\mathbf{r})$.

This is illustrated in Fig. 1(b), which shows the way in which the temperature varies in a thin flat metal plate in which there is a heat source at one point. The plate has its upper boundary running along the top of the picture, and the point heat source is a distance a below the boundary. Surrounding the heat source we see contours of constant temperature (labeled by T=const); the temperature decreases as we move away from the heat source. Note that because the heat cannot flow across the boundary of the plate, it must flow along the edge of the plate away from the source. We also plot some streamlines showing the flow of HEAT (labelled by K). The heat always flows from high T to low T, which is why the heat flow streamlines are everywhere perpendicular to the temperature contours. Actually the streamlines of heat flow show the 'vector field' of heat flow - see below for what a vector field is.

Thus we see that we can define a 'temperature field' $T(\mathbf{r})$ in the plate, which varies continuously with position $\mathbf{r}$ around the plate.

(2) VECTOR FIELDS: Now let's consider a situation where the quantity that is varying is not a number, but a *vector*. A vector is simply a quantity which has both a magnitude and a direction. For example, consider the flow of some current (eg., a current of air, or water); at every point $\mathbf{r}$ in space, this current is defined by the velocity $\mathbf{v}(\mathbf{r})$ with which the current is flowing in some specified direction. Thus we get a 'velocity field' $\mathbf{v}(\mathbf{r})$ defined throughout the domain. Another familiar example is provided by a *Force*; recall that a force has a magnitude and direction. Thus, eg., we can think of the 'gravitational force field' around a massive body; this is just a map of the force vectors $\mathbf{F}(\mathbf{r})$ at every point $\mathbf{r}$ around the body.

Let's consider how we might make such a map (see Fig 1(c)). Imagine mapping out the current in some stream at a whole variety of different points, by measuring the velocity at these points. We might end up after a while with the result shown in the top left part of Fig 1(c). Each vector is shown as an arrow - the direction of the arrow shows the direction of the current, and the length of the arrow shows its magnitude. Now imagine trying to make a continuous map out of this - we will end up with something like the 2nd picture at bottom right of Fig. 1(c). We see that we can join the arrows together to make streamlines, showing the current flow throughout the domain. As a matter of fact, the magnitude of the velocity at some point on the streamline is just proportional to the density of streamlines - if they are closely bunched together, this means a higher velocity.

We now see how the heat current $\mathbf{K}(\mathbf{r})$ in Fig. 1(b) is a kind of vector field - for the heat flow is just another vector, with a direction and a magnitude. In the same way let us plot the 'electric field' lines around a positive charge Q. As we will see, these field lines at some point $\mathbf{r}$ near the charge are defined by the force exerted by the charge on some infinitesimal 'test charge' situated at $\mathbf{r}$. We see that the field lines show a 'radial' outward pattern - they are directed everywhere away from the charge. If we were dealing with the gravitational field from a point-like mass M, we would see that same radial pattern, but it would be directed inwards (gravity attracts).

Note, incidentally, that we can always associate a vector field with any scalar field. We saw this with the temperature field - the vector field associated with this describes the *gradient*, or 'slope' of the scalar temperature field - it has a direction perpendicular to the contour lines (ie., along the direction which takes us from higher to lower temperature), and magnitude equal to the rate at which the temperature is changing along this direction. The words 'gradient' and slope' come from the vector field associated with a an altitude contour map. If we draw 'streamlines' perpendicular to the lines of constant altitude on a topographic map, we just get lines showing us the direction of the slope or 'gradient' of the topography - and the magnitude of these vectors (ie., the density of the gradient lines) shows how steep the slope is.

You might ask whether we can define even more complicated fields in some domain, and whether these are physically significant. The answer is yes, we can. In fact we can imagine that at each point there may be a whole bunch of vectors, representing some complicated physical variable. These fields are called 'tensor fields'. In general these are too complicated to discuss here, but we will see that they are nicely exemplified in things like the 'stress' in a solid, or the 'curvature' in some medium.

Finally, notice that all of the fields that are given above as examples are imagined to exist in space. However, we also notice that all of these fields can in principle change with time. Thus, the temperature will change with time, in different ways at different points in space; likewise the current flow in a stream, or the heat flow in a room (indeed, if the temperature changes, so must the heat flow, since they are related). Even the topography of some landscape will change in time (albeit rather slowly on human timescales). For a mathematical definition of the field this entails no real change - we simply extend the domain of the field to include time as well as space, and specify the 'position' of the field in this domain as its position in both time and space. Thus, eg., instead of talking about the temperature T as a function only of $\mathbf{r}$, we now have it depend also on the time t, so we have a scalar field $T(\mathbf{r}, t)$ varying in space and time.

FIELDS in the PHYSICAL WORLD

It is of course fine to cook up models of the real world in terms of continuous fields defined in space and time. But to what extent do these models correspond to phenomena in the real world? This is a question about physics, not mathematics, and the standard way of answering this question in modern science is to see how well the predictions of such models correspond to the behaviour of real world systems.

The first objection one might have is that in the real world, things may not vary continuously. Thus, even of one could accept that space and time are something continuous, a little thought shows that a quantity like fluid velocity or temperature might not be a continuous quantity. After all, if we accept that matter is made of atoms, and if it is supposed that matter cannot exist at a smaller scale than atoms, then it seems meaningless to talk about things varying continuously at the scale of atoms. After all, the argument would go, at a particular point in space, one either has an atom or not - there is no continuous transition between these two states!

If one assumes a picture of atoms as indivisible/irreducible object, such a question clearly poses a problem for a continuous field description. However current views on physics assume that it is meaningful to talk about 'parts' of objects, at least down to length scales and timescales which are incredibly small (down to the 'Planck length' and 'Planck time'). And, since we have no idea how things are supposed to behave at these scales, it is not inconsistent to imagine going to even smaller scales. Under these circumstances, a continuous description of atoms and parts of atoms is perfectly meaningful. However one should always bear in mind the possibility that Nature may in some sense be discontinuous (a possibility that poses its own problems - for if there is, eg., a smallest unit of matter, then one immediately is led to ask - what is it made of?).

The idea that the world might be made of some fundamental elements or types of stuff is of course very old (certainly older than the pre-Socratic philosophers, where we first meet it discussed in detail). However a radically new idea began to emerge when physicists unraveled the nature of electromagnetic phenomena, in the work of Faraday and Maxwell. This was the idea that the fundamental building blocks of Nature were not particles, or atoms, or some similar set of objects localized in space. Instead, the fundamental entities were delocalized *fields*, spread throughout space and time. As such they have a much more elusive (one might say aetherial) nature - indeed, they are not visible to us directly at all, but only through their effects. We cannot directly experience an EM field, but only observe its effects on electric charges, which happen to be tied to matter. This fact requires us to develop a new idea, that of the 'test charge'. Since the only way that one can detect the presence of some field is by its effect on some test charge (this charge being associated with that particular field), one needs to imagine that we 'decorate' the entire domain of the field with test charges, and observe their motion, caused by the field. There is a catch here, however - the same charges that react to the field by moving also create a polarization of the field, ie., they distort it - so we need to imagine that the test charges are very small, in fact infinitesimal. In the case of EM fields, the test charges are typically electrons; in the case of a gravitational field, they are tiny masses (we can imagine a very rarified dust, distributed in space around a large mass). In either case, we can only become aware of the existence of the field though its effect on these test charges.

From a purely philosophical point of view, this result, that the fundamental building blocks of Nature are not immediately present to us, should be considered as an asset. As stressed in the earlier discussion of Greek philosophy, one of the great advances made by the Greeks was to recognize that any fundamental description of Nature had to deal with entities which were *not* immediately available to us - the world presents us with a bewildering array of phenomena and materials, and any explanation of these must inevitably reach to some more fundamental (and presumably rather different) objects which *underlie* the world of surface appearances. When we come to discuss later philosophy (notably that of Kant) this point will become much clearer.

It is tempting to think of the EM field as being the realization of the long-sought aether. It conforms rather closely with Huyghens's idea of the aether as the medium in which EM waves are transmitted - one can think of the EM field as having properties similar to a very stiff medium through which waves can move at very high velocity. This medium, in EM theory, only interacts with electrical charges, not with electrically neutral matter - this is why an object like a planet (or any other neutral object) moves freely though it. This of course was the key piece of information that was unavailable to Newton and Huyghens - the existence of other forms of matter (ie., electric charge) than electrically neutral mass, which could interact with an aether and also with matter. Thus it appears at first glance that one can have a consistent theory in which one has neutral matter, plus electric charges, plus the EM field, with this last behaving like the aether. However, as noted above, a basic problem arises, first uncovered fully by Einstein. This is that if an aether exists in spacetime, it must 'be' somewhere - ie., it is like a medium which is stationary in some spatial reference frame. However it turns out that the actual behaviour of EM fields is incompatible with this idea - one sees that the velocity of waves in the EM field is observed to be the exactly same no matter what the observer might be doing (thus, 2 observers, moving rapidly with respect to each other, would still see a light beam moving at the same velocity). This makes no sense if light is thought to be moving through some medium at a given velocity.

Curiously, however, the idea of an aether received some stimulus from the General Theory of Relativity. In this theory, spacetime itself can be thought of as a field (for a proper discussion see the notes on this topic). This field can itself support waves (called 'gravity waves') just as though it was a medium of some kind; and like the EM field, it interacts with a kind of charge, which is in fact just mass/energy. In the same way as the EM field, spacetime can be distorted or 'polarized'; in the EM field, such distortions correspond to magnetic or electric fields, whereas in the case of spacetime, the distortions of spacetime are what we call gravitational fields. In EM fields, the magnetic or electric distortions are actually caused by charges themselves and their motion, and they then act back on the charges to change their motion. Likewise, the distortions of spacetime that correspond to gravity are caused by mass/energy, and these distortions then act back on the mass/energy, changing its motion. Thus one can at least argue that the idea of the aether has returned in the form of a field, albeit in a rather stranger form than originally imagined - in the form of different kinds of field.

The subsequent results of 20th century physics have only tended to confirm this idea of the field as a fundamental entity in Nature. It is now clear that we can think of all of matter as being composed of various kinds of matter fields called 'fermionic fields', which interact with each other via other kinds of fields called 'bosonic fields'. These fields are not simple and are moreover quantum-mechanical; but the theory of this, which was finally completed in the 1970's under the name of the 'standard model' (accomplished notably by the theorists Salam, Weinberg, 't Hooft, Wilczek and Gross), has been enormously successful and has withstood all attempts to falsify it. The latest test of it, in 2012, was passed with the successful observation at CERN of the 'Higgs boson', one of the many different subfields embodied in the standard model.

However there is still a very large 'fly in the ointment' here. The General Theory of Relativity has been just as successful as the Standard Model in its explanation and prediction of physical phenomena. As noted above, this theory is also a field theory in which spacetime itself a field. However, General Relativity has resisted all attempts to make sense of it in quantum mechanical terms. While the long-running programme of research that is called 'string theory' has been attempting to do this since the 1970's, one can also adopt the point of view that quantum mechanics and General Relativity are fundamentally incompatible, and that one or both of them must eventually fail.

Thus we are currently faced with a rather peculiar situation - we have 2 successful theories (the 'quantum theory of fields', and General Relativity), which may be fundamentally incompatible. Both of them are extremely successful, and we have yet to find a situation in which either of them fails, which we can also check out experimentally. Perhaps the greatest challenge in physics is to somehow confront these two theories, and see what happens. Very few physicists would claim to have much of an idea of what will come out of such a confrontation. It will be interesting to see if the field concept, as we presently understand it, will survive unscathed.