LAWS OF NATURE:

CREATION AS REDUCTION OF SYMMETRY

E.C.G. Sudarshan

Indian Institute of Science

Bangalore, India

and

University of Texas

Austin, Texas

In ancient India we had the notion of a level of functioning beyond kingship, that of the "ćakravarti" (the Maintainer of the Wheel) who was the embodiment of the supremacy of the functioning. At this level all external norms are absent and the ćakravarti evolves the norms from within: and functions in accordance with natural law.

Our present understanding of the physical universe owes a great deal to the way of approaching dynamics as set out by Galileo and Newton. The latter enumerated a set of "laws of motion" which introduce the notions of inertia and impressed forces, the quantitative relation between change of momentum (quantity of motion) and the impressed force and enumerated the law of conservation of the total momentum of a system whose parts are in interaction. We now know that these three laws of motion are all of different categories. The First Law ("A body continues in its state of rest or of uniform motion unless acted upon by external force"), comprising in itself a bold declaration about the nature of inertia (that uniform motion needed no cause!), defines the free particle. The Second Law ("The rate of change of momentum is equal to the impressed force") is both a definition of the concept of the force and a quantitative measure of the same. The Third Law ("Action and reaction are equal and opposite") is a relationship between forces exerted by parts

of a system on each other and says that the total momentum of an isolated system is conserved notwithstanding any mutual interaction; it is the forerunner of a series of conservation laws which derive from invariance--in this case the observation that all locations in space are equivalent for an isolated system and that as a consequence the total momentum is a conserved quantity.

These laws set a pattern for dynamics. The inertia of a "particle" (the prototype isolated system) is an intrinsic quality of the particle. When there is a collection of mutually interacting particles the simpler systems can be characterized by a potential energy function from which the mutual forces can be derived. This way of approaching the problem has come to be known as Lagrangian mechanics. In this generalization the concept of motion is liberalized: instead of looking at a point (or collection of points) moving in time we may have "generalized coordinates" evolving in time. Such an evolution produces the trajectory of the system. Physics is thus poised for describing general dynamical evolution of systems richer than moving particles.

The next generalization is known as Hamiltonian dynamics. Here position and momentum (or more generally, configuration and flow) are treated on par to describe the "phase" of the system. Dynamics then describes the evolution of the phase, the phase space trajectory. The dynamical content is still

the same, the force or potential now being generalized into a Hamiltonian. The realization that uniform motion is an aspect of inertia is however realized in treating position and momentum on the same footing.

Within the Hamiltonian scheme an entirely independent view of dynamics becomes most transparent. We have already talked about the Third Law of Motion being related to the irrevelance of the origin of coordinates chosen to describe the motion. This is one of the "invariance" properties of Newtonian dynamics. There are other such requirements that we wish to impose on the dynamics of an isolated system. The irrelevance of the orientation of the coordinate axes implies the constancy of the angular momentum (the moment of momentum); and the irrelevance of the absolute origin of time implies the constancy of the energy. Thus the invariance properties are related to conservation laws: irrelevance of some quantities implies the immutability of some (conjugate) quantities!

One can go one step beyond this to classify and even crystallize dynamical laws which admit of certain invariance principles.

We might go further and ask for the simplest realizations of the invariance group. To our pleasant surprise we see the (Newtonian) particle emerge in this analysis as the simplest possible dynamical system: the so-called irreducible

realization. The mass appears as an essential label of the realization. So the roles of the particle and of the invariance group are now interchanged: the particle becomes the crystallization of the invariance properties and is implicitly contained in it rather than the particle with its dynamical potentialities exhibiting the invariance properties. Newton's First Law in a higher and richer setting provides the particle as its realization.

A great advantage of this method of deriving a particle from the symmetry principles is that the method is equally applicable to quantum theory. The quantum particle can be cognized as the simplest realization of the invariance principles of dynamics and dynamical systems can all be built up from such irreducible components. The "word becomes flesh": śabda entails dravya. (In this context one is reminded of the stanza:

> "ṛṣaya pitaro deva sarva bhūtāni dhātava
> jangamājangamam ćedam jagannārayanodbhavam").

One may also note that in this view of the genesis of dynamical systems there are no causes and effects, no agents, no agencies. There are not even intrinsic differences between various states of motion: all differences are differences entailed by frame choices. What is fixed in one frame is moving in another. Uniform motion is uncaused. Time is the

vehicle for unfolding: the anterior and posterior configurations are related but there is a symmetric relation between them, not a directed cause-effect relation. Collections of such particles constitute the physical universe.

Can such ideas be relevant for interacting systems? Can we derive the forces and the changes of motion from a few principles of symmetry? Is there the possibility that there is a grand unified picture of dynamics? In other words can we perceive dynamical law as natural law?

We are all aware of the heroic deeds of Sir Isaac Newton who saw in the falling apple the operation of the force of gravity; and saw that too as an aspect of the law of universal gravitation which manifested itself in such diverse phenomena as the motion of tides, the orbit of the moon around the earth and of the planets around the sun: Many systems, many motions, but one law.

But one law is still a law: there is now a cause for motion changes, namely gravitation. Who or what imposed this force, this cause for changes of momentum? Have we broken out of the "causeless evolution" characterizing collections of free particles? Not necessarily, if we can ascend to a more natural method of describing the natural evolution of a system. Since the gravitational force according to Newton is proportional to the mass of the particle on which it acts, it follows that the acceleration caused by the gravi-

tational force is independent of the mass: consequently every particle which is subject to gravitation has an acceleration which depends on the location: we have a modification of space itself. Albert Einstein used this as the fundamental starting point: under the influence of gravitation space itself undergoes a change in which uniform rectilinear motion is impossible. Space becomes curved. Gravitation is the same as curvature of space. Einstein demanded further that the manner in which the curvature comes about must (1) lead to the Newtonian gravitational equation for weak gravitational fields; (2) it must incorporate the principle of relativity; and (3) the description must be unchanged by the specific choice of coordinates in a curved space. These principles led him to a theory of gravitation in terms of the curvature of space-time and the curvature itself being related to the energy and momentum distribution in space.

This theory had a surprising consequence. Even in planetary motion and stellar astronomy the Newtonian predictions had small but computable deviations from the new Einsteinian predictions; and the latter were in agreement with measurements. Emboldened by this we looked for more general applications including the evolution of space-time curvature itself. This led to an understanding of the "expansion of the universe" and more generally served to make cosmology a realm of physical science.

In a uniform space we would expect to define measures of lengths which are "the same" at each point: we would expect that if the unit of length is defined at one location we can "translate" that unit of length. In a curved space this is given by specifying a "connection," a rule to specify how "sameness" is to be realized. Such a rule is necessary since any method of assigning rectilinear coordinates in a curved space is valid only over a limited region (a "patch"). The curved space is mathematically specified by these coordinate patches which collectively cover the entire space-time and the connections. While the coordinate labels and the numerical quantities that specify the connections depend on the scales and orientations of the rectilinear coordinate systems in each patch (the "gauge") there are quantities that describe the space which are invariant under such changes. These "gauge-invariant" quantities should be used to describe the physics of gravitation: they are the curvature tensors. Einstein found a relationship between curvature of space and the distribution of energy and momentum in space which supplants Newtonian description. Gravitation is thus made a natural law--it is made to inhere in the fabric of space-time.

An immediate bonus of such an approach is the ability to do physical cosmology: to ask questions about this very universe, at least in its large-scale properties. The non-static solutions of the Einstein equations for cosmology

produce for us a model of an expanding universe, quite in accordance with the discovery of red-shifts of galaxies and a mutual recession of all galaxies with a speed proportional to their mutual separation. The evolution of the existing universe itself is subject to a natural law.

The gauge principle which states that "interaction" is a consequence of the requirement of local symmetry transformations is of much wider applicability than gravitation. In a "field theory" the dynamical configurations are specified by the assignment of a collection of fields at each space-time point. Field equations describing the evolution connect the space-time derivatives and the fields, and are usually differential equations. Symmetries of the fields imply a multiplet of fields at each point which can be replaced by linear combinations without changing the equation of motion. Just as the equations of motion of Newtonian particles in empty space are unchanged by change of coordinate systems, the internal symmetries also mean independence of the internal coordinate system. A local symmetry means that you can change the internal coordinate system from one space-time point to another. Since differential equations involve derivatives which are rates of changes of fields with space and time, these local symmetries can hold only if there is an associated set of gauge fields coupled to it in a specific manner. Thus

a gauge principle can introduce natural interactions. The unification of electric and magnetic interactions by Maxwell into the theory of electromagnetism may be viewed as the natural gauge symmetry-induced interaction where the symmetry is associated with the electric charge. The electromagnetic potentials are coupled to the density and flow of electric charge in this theory, somewhat like the gravitational potentials were coupled to the density and flow of energy-momentum.

In both gravitation and electromagnetism we have long-range forces; the familiar inverse square law forces are an expression of this feature. It does appear that these are the only long-range forces in nature. The interactions responsible for holding the nucleus together (the "strong interactions") and the interactions responsible for radioactive transmutations (the "weak interactions") seem to be definitely short-range interactions. While the strong interactions appeared to be rather complicated the weak interactions are really quite simple. Twenty-five years ago Sudarshan and Marshak showed that they are very similar to electromagnetism: the coupling involved a flow of only left-handed objects. The weak interaction is therefore called "chiral." About fifteen years ago Weinberg, Salam and Glashow discovered that the Sudarshan-Marshak interactions too could be seen as gauge interactions of which Maxwell

electromagnetism is a component part. So this has come to be called the "unified electro-weak theory."

As a natural sequence to this it has been discovered that the strong interactions could be seen as a gauge theory associated with yet another kind of internal symmetry called "color" so that the corresponding field theory is called "chromodynamics." In this theory the nuclear particles are composed of yet more elementary objects called "quarks" which have and can change color. It is the coupling to the flow of color that is generated by local color invariance that lead to all the strong interactions as well as the composite structure of the nuclear particles.

The possibility of accounting for short-range forces using gauge theories suggests the possibility of unifying electro-weak and strong interactions into what is called a grand unified theory.

(cf. bahūnām janmanāmante
j̃nānavān mām prapadyate
vāsu deva sarvamiti
sa mahātma sudurlabhaḥ.)

Such theories are now being actively studied on the basis of their great aesthetic appeal and the possibility of new phenomena that could be explained.

The mechanism of generating short-range interactions using gauge theory makes use of spontaneous symmetry breaking. This is a fascinating phenomenon in which one is working at the edge of causality.

We are familiar with many cases in which the state of lowest energy (the "ground state") is fully symmetric under the symmetries of the problem. If we drop a ballbearing into a circular bowl, the ballbearing will seek the lowest energy when it is at rest in the centre of the bowl. This is a unique configuration, and it looks the same from all horizontal directions. It is rotationally invariant. On the other hand if the bowl had a circular blister in the centre the lowest energy would not be reached at the centre: the centre would not be a location of stable equilibrium. In its place we have a multitude of equilibrium locations, none of which exhibits the circular symmetry of the bowl. A simpler situation is obtained by a ballbearing at rest on a horizontal table. All positions are equally good but none of them exhibits the simple translation symmetry of the table top.

In such a case no choice has the symmetry of the problem and <u>any</u> choice breaks the symmetry. The symmetry is broken without cause, broken spontaneously. Motion about the equilibrium position, at least along certain directions becomes very easy: as easy as rolling the ballbearing on the table. As soon as this situation develops in a field theory the

interaction becomes a short-range interaction, the inverse of the range being proportional to the departure of the equilibrium state from symmetry. A large symmetry breaking for the ground state implies a short range for the interactions.

It follows that if we started with a large set of symmetries some of which are spontaneously broken, then some of the interactions would be short range and some long range even though they are all gauge interactions. The possibility of spontaneous symmetry breaking thus generates richer patterns. We recognize that the structure and functioning of the world depends crucially on the short-range nature of the nuclear and weak interactions contrasted with the long-range electromagnetic effects: the latter (together with the underlying quantum framework) provide a flexible electronic envelope of the atom which accounts for most of the chemical, optical and structural properties of matter. This could not be so if all the interactions were long range or all of them short range. Thus in a curious way the symmetry of an ice crystal or a diamond crystal depends on the spontaneous breaking of a deeper symmetry. The beauty of the world, the very act of creation, involves the breaking of symmetry. Creation is the reduction of symmetry.

It is beyond the scope of this article to discuss the ramifications of these modern concepts to cosmology and particle physics. It is however important to observe that

natural philosophy has now reached a new depth in probing nature. Nature's laws themselves are the data for providing a deeper theory. In the search for the cause for the diversity in the universe we are led to the causeless breaking of symmetry. The potentialities are fully describable but the realization of a broken symmetry world is due to the instability of the fully symmetric configuration.

We may make a general comment about natural law. We expect that at a deep enough level, in place of a number of different laws, we expect to find a unified law. However any scheme for unification must contain within it the mechanism for the universality to be broken. In other words, both the symmetry and its breaking must be aspects of the formulation of natural law.