PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A
MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

Dimensionality

J. D. Barrow

Phil. Trans. R. Soc. Lond. A 1983 **310**, 337-346
doi: 10.1098/rsta.1983.0095

Phil. Trans. R. Soc. Lond. A **310**, 337–346 (1983) **337**
Printed in Great Britain

Dimensionality

By J. D. Barrow
*Astronomy Centre, University of Sussex, Brighton BN*1 9*QH, U.K.*

We examine the role played by the dimensions of space and space–time in determining the form of various physical laws and constants of Nature. Low dimensional manifolds are also seen to possess special mathematical properties. The concept of fractal dimension is introduced and we discuss the recent renaissance of Kaluza–Klein theories obtained by dimensional reduction from higher dimensional gravity or supergravity theories. A formulation of the anthropic principle is suggested.

1. Historical introduction

The fact that we perceive the world to have three spatial dimensions is something so familiar to our experience of it that we seldom pause to consider the influence this special property has upon the laws of physics. Yet, it appears that the dimensionality of the world plays a key part in determining the form of the laws of physics and in fashioning the roles played by the constants of Nature.

Interest in explaining why the world has three dimensions is by no means new. From the commentary of Simplicius and Eustratius (see Neugabauer 1975), Ptolemy is known to have written a study on the 3D nature of space entitled *On dimensionality* in which he argued that no more than three spatial dimensions are possible in Nature. Unfortunately this work has not survived. What does survive is evidence that the dramatic difference between systems identical in every respect but spatial dimension was discovered and appreciated by the early Greeks. The Platonic solids, first discovered by Theaitetos (Sarton 1959), brought them face-to-face with a dilemma: why are there an infinite number of regular, convex, two-dimensional *polygons* but only five regular three-dimensional *polyhedra*? This mysterious property of physical space later spawned many mystical and metaphysical 'interpretations'.

In the modern period, mathematicians did not become actively involved in attempting a rigorous formulation of the concept of dimension until the early nineteenth century. During the nineteenth century Möbius considered the problem of superimposing two enantiomorphic solids by a rotation through 4-space and later, Cayley, Riemann and others developed the systematic study of N-dimensional geometry although the notion of dimension they employed was entirely intuitive. It sufficed for them to regard dimension as the number of independent pieces of information required for a unique specification of a point in some coordinate system. Gradually the need for something more precise was impressed upon mathematicians by a series of counter-examples and pathologies to several simple intuitive notions. For example, Cantor and Peano provided injective and continuous mappings of $\mathbf{R}$ into $\mathbf{R}^2$ to confute ideas that the unit square contained more points than the unit line. After unsuccessful attempts by Poincaré it was Brouwer who established the key result (Brouwer 1911; 1913): he showed that there is no continuous injective mapping of $\mathbf{R}^N$ into $\mathbf{R}^M$ if $N \neq M$. The modern definition of dimension due to Menger and Urysoln grew out of this fundamental idea (Menger 1928; Hurewicz & Wallman 1941).

The question of the *physical* relevance of spatial dimension seems to arise first in the early work

(1747) of Immanuel Kant (see Handyside 1929). He realized that there was an intimate connection between the inverse square law of gravitation and the existence of precisely three spatial dimensions, although he regards the three space dimensions as a consequence of Newton's inverse square law rather than vice versa.

In the twentieth century a number of outstanding physicists have sought to accumulate evidence for the unique character of physics in three dimensions. Ehrenfest's famous article of 1917 was entitled *In what way does it become manifest in the fundamental laws of physics that space has three dimensions?* and it explained how the existence of stable planetary orbits, the stability of atoms and molecules, the unique properties of wave operators and axial vector quantities are all essential manifestations of the dimensionality of space. Soon afterwards Herman Weyl (1922) pointed out that only in $(3+1)$-dimensional space–times can Maxwell's theory be founded upon an invariant integral form of the action; so, only in $(3+1)$ dimensions is it conformally invariant.

In recent years the problem of dimensionality has re-emerged in connection with supersymmetric gauge theories and the description of chaotic dynamical systems. Our aim here is to give a brief discussion of some of the most interesting influences of dimensionality on physics and mathematics.

2. Orbits

One of the first areas of physics to display the role of dimensions is the theory of orbital motion under central forces. We immediately see the reason for the prevalence of inverse square laws in physics. Consider first the question of planetary motions.

The Poisson–Laplace equation for the gravitational field of force in an N-dimensional space has a solution for the gravitational potential, ϕ and force, F, of the form

$$\phi(r) \propto r^{2-N}; \quad F(r) \propto r^{1-N}, \quad N > 2, \tag{1}$$

for a radial distribution of material. The inverse square law of Newton follows as an immediate consequence of the tri-dimensionality. A planetary motion can only describe a central elliptic orbit in a space with $N = 3$ if its path is *circular*, but such a configuration is unstable to small perturbations. In three dimensions, of course, stable *elliptical* orbits are possible. If hundreds of millions of years in stable orbit around the Sun are necessary for planetary life to develop then such life might only develop in a three-dimensional world. In general, the existence of stable, periodic orbits requires $r^3F(r) \to 0$ as $r \to 0$ and $r^3F(r) \to \infty$ as $r \to \infty$. Therefore we require $N < 4$. If we examine the analogous problem in general relativity by examining motion in the gravitational field of the $(N+1)$ dimensional Schwarzschild solution we again find no stable bound orbits exist for $N > 3$.

One of Newton's most famous results is his proof that if two spheres attract each other under an inverse square law of force they may both be replaced by points concentrated at the centre of each sphere with mass equal to that of the respective sphere. It can be shown (Sneddon & Thornhill 1949) that the general form of the gravitational potential for which the gravitational force of a sphere can be replaced by that of a point at its centre is the Yukawa potential

$$\phi \propto e^{-\lambda r}/r, \tag{2}$$

although, for a sphere of radius a and density ρ, the mass that must be concentrated at its centre is

$$M(\lambda) = (4\pi a\rho/\lambda^2)\,(\cosh \lambda a - (\sinh \lambda a)/\lambda a), \tag{3}$$

and we recover Newton's result as $\lambda \to 0$

$$M(0) = \tfrac{4}{3}\pi a^3 \rho. \quad (4)$$

The truth of Newton's result is a direct consequence of the existence of three spatial dimensions in (1). Gravitation physics is simplest in three dimensions.

3. Atomic stability

It is widely known that matter is stable. By this we mean that the ground state energy of an atom is finite. However, the common text-book argument that employs the uncertainty principle to demonstrate this is actually false. Although the energy equation for a single electron of mass m and charge $-e$ in circular orbit around a charge $+e$ gives a total atomic energy of

$$E = h^2/2mr^2 - e^2/r, \quad (5)$$

and this energy apparently has a finite minimum of $r_0 \approx h^2/2me^2$, where $E'(r_0) = 0$, it is, in principle, possible for the electron to be distributed in a number of widely-separated wave packets. The packet close to the nucleus could then have an arbitrarily sharp momentum and position specification at the expense of huge uncertainty in the other packets. In this manner the ground state energy might be made arbitrarily negative.† For these reasons analyses of atomic stability such as those of Ehrenfest (1917) and those that use only the uncertainty principle must be regarded as only heuristic. However, their results are confirmed by an exact analysis of the Schrödinger equation. In 1917, Ehrenfest considered only the simple Bohr theory of an N-dimensional hydrogen atom. He found the energy and radii of the energy levels and noted that when $N > 5$ the energy levels increase with quantum number whereas the radii of the Bohr orbits

$$r_L(N) \approx (me^2 L^{-2} h^{-2})^{1/(N-4)}$$

decrease with increasing quantum number L and electrons just fall into the nucleus. Alternatively, if we write down the total energy for the system and use the uncertainty principle to estimate the kinetic energy resisting localization we have for $N > 2$, that

$$E \approx h^2/2mr^2 - e^2/r^{N-2}. \quad (6)$$

It can be seen that for $N > 5$ there is no energy minimum. For $N = 4$ the situation is ambiguous because there ceases to exist any characteristic length in the system. This also indicates that no minimum energy scale can exist. It is possible to demonstrate this more rigorously by including special relativistic effects in the energy equation (6). Thus, for $N = 4$, the relativistic energy is (where m_0 is the rest mass of the electron now),

$$E \approx (c^2h^2/r^2 + m_0^2c^4)^{\frac{1}{2}} - e^2/r^2, \quad (7)$$

and so as $r \to 0$, $E \to -1/r^2$ and no stable minimum can exist.

On the basis of these arguments it could be claimed that, if we assume the structure of the laws of physics to be independent of the dimension, stable atoms, chemistry and life can only exist in $N < 4$ dimensions. (Note that in two dimensions all energy levels are discrete and there exists a finite energy minimum together with a spectrum extending to infinity, the radius of the first orbit is huge, *ca.* 0.5 cm.) These simple arguments can be confirmed by solving the Schrödinger equation for the N-dimensional hydrogen atom. The dimensionality of the Universe is a reason for the existence of chemistry and therefore, most probably, for chemists also.

† A much stronger, *nonlinear* constraint is required in addition to the Heisenberg uncertainty principle if one is to rule out ground state energies becoming arbitrarily negative. The strongest result is supplied by the non-linear Sobelev inequality (Lieb 1976). This supplies the required bound on the ground state energy and shows that matter is indeed stable in quantum theory.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

4. Wave equations

Many authors have drawn attention to the fact that the properties of wave equations are very strongly dependent upon the spatial dimension (Ehrenfest 1917; Poincaré 1917; Hadamard 1923). Three-dimensional worlds appear to possess a unique combination of properties which enable information-processing and signal transmission to occur via electromagnetic wave phenomena. Since our Universe appears governed by the propagation of classical and quantum waves it is interesting to elucidate the nature of this connection with dimensionality.

Let us recall, as simple examples, the solutions to the simple classical wave equation in one, two and three dimensions.

One dimension:

$$\frac{1}{c^2}\frac{\partial^2 u}{\partial t^2} = \frac{\partial^2 u}{\partial x^2}, \tag{8}$$

where c is the signal propagation speed,
with initial conditions set at $t = 0$ as

$$\left.\begin{aligned} u(x,0) &= f(x),\\ \frac{\partial u}{\partial t}(x,0) &= g(x). \end{aligned}\right\} \tag{9}$$

This has the solution of D'Alembert,

$$u(x,t) = \tfrac{1}{2}[f(x+ct)+f(x-ct)] + \frac{1}{2c}\int_{x-ct}^{x+ct} g(y)\,\mathrm{d}y. \tag{10}$$

Two dimensions:

$$\frac{1}{c^2}\frac{\partial^2 u}{\partial t^2} = \frac{\partial^2 u}{\partial x^2}+\frac{\partial^2 u}{\partial y^2}, \tag{11}$$

with initial conditions at $t = 0$ of

$$\left.\begin{aligned} u(x,y,0) &= f(x,y),\\ \frac{\partial u}{\partial t}(x,y,0) &= g(x,y). \end{aligned}\right\} \tag{12}$$

This has the solution of Poisson,

$$u(x,y,t) = \frac{1}{2\pi c}\frac{\partial}{\partial t}\iint_{\rho\leqslant ct}\frac{f(\xi,\eta)}{(c^2t^2-\rho^2)^{\frac{1}{2}}}\,\mathrm{d}\xi\,\mathrm{d}\eta + \frac{1}{2\pi c}\iint_{\rho\leqslant ct}\frac{g(\xi,\eta)}{(c^2t^2-\rho^2)^{\frac{1}{2}}}\,\mathrm{d}\xi\,\mathrm{d}\eta. \tag{13}$$

$$\equiv \rho^2 = (\xi-x)^2+(\eta-y)^2.$$

Three dimensions:

$$\frac{1}{c^2}\frac{\partial^2 u}{\partial t^2} = \frac{\partial^2 u}{\partial x^2}+\frac{\partial^2 u}{\partial y^2}+\frac{\partial^2 u}{\partial z^2}, \tag{14}$$

with initial conditions at $t = 0$ as

$$\left.\begin{aligned} u(x,y,z,0) &= f(x,y,z),\\ \frac{\partial u}{\partial t}(x,y,z,0) &= g(x,y,z). \end{aligned}\right\} \tag{15}$$

This has the solution of Kirchoff,

$$u(x,y,z,t) = \frac{1}{4\pi c^2}\frac{\partial}{\partial t}\left[\frac{1}{t}\iint_{r=ct} f(\xi,\eta,\zeta)\,\mathrm{d}S\right] + \frac{1}{4\pi c^2 t}\iint_{r=ct} g(\xi,\eta,\zeta)\,\mathrm{d}S, \tag{16}$$

where $r^2 = (\xi-x)^2+(\eta-y)^2+(\zeta-z)^2$ and $\mathrm{d}S$ is the surface element with respect to (ξ,η,ζ) on the sphere $r = ct$ centred on $(x,y,z) = (0,0,0)$.

From these three solutions, (8)–(16), something remarkable emerges. We see that in the one- and two-dimensional cases, the domain of dependence that determines the solution $u(x,t)$ at point (x,t) is given by the closed interval $[x-ct, x+ct]$ and the disk (interior plus boundary) $r \leqslant ct$, respectively. Therefore, in both cases the signals may propagate at any speed less than or equal to c. In complete contrast, the three-dimensional solution has a domain of dependence consisting only of the *surface* of the sphere of radius ct. All three-dimensional wave phenomena travel *only* at the wave velocity c. We are ignoring the effect of dispersion here.

What this means in practice is that in two-dimensional spaces signals emitted at different times can be received simultaneously: signal reverberation occurs. It is impossible to transmit *sharply* defined signals in two dimensions, for example, by waves on a liquid surface. Now it is known that the transmission of wave impulses in a reverberation-free fashion is impossible in spaces with an *even* number of dimensions (Hadamard 1923). The favourable odd dimensional cases are said to obey Huygen's Principle (Courant & Hilbert 1962). This situation has led many to suppose that life could only exist in an odd dimensional world because living organisms require high fidelity information transmission at a neurological or mechanical level (Poincaré 1917; Whitrow 1955).

Interestingly, one can narrow down the number of reasonable odd dimensional spaces even more dramatically by appealing to the need for wave signals to propagate *without distortion*. Three-dimensional worlds allow spherical waves of the form

$$u(x_1, x_2, x_3; t) = h(r) f(r-ct), \tag{17}$$

with

$$r^2 = \sum_{i=1}^{N} x_i^2, \tag{18}$$

to propagate in *distortionless* fashion to large distances; but this is no longer the case for $N > 3$. For example, in seven dimensions Ehrenfest shows that a solution to the wave equation is,

$$u(x_1, \ldots, x_7; t) = (A/r^5) f(t-r/c) + (B/r^4) f'(t-r/c) + (D/r^3) f''(t-r/c), \tag{19}$$

where A, B and D are constants. Thus, at time t there is no reverberation; only signals which were emitted at the time $(t-r/c)$ are received. However, these signals are now strongly distorted because at large r the terms in f'' and f' determine the form of $u(\boldsymbol{x}, t)$.

Only three-dimensional worlds appear to possess the 'nice' properties necessary for the transmission of high fidelity signals because of the simultaneous realization of reverberationless and distortionless propagation.

5. Fundamental units

Our Universe appears to possess a collection of fundamental or 'natural' units of mass, length and time that can be constructed from the physical constants G, h and c, (see Barrow 1983). A dimensionless constant can only be constructed if the electron charge e is also admitted and then we obtain the dimensionless quantity e^2/hc first found by Sommerfeld. In a world with N dimensions the units of h and c remain ML^2T^{-1} and LT^{-1} but the law of gravitation changes in accord with (1) and hence the units of G are $M^{-1}L^N T^{-2}$. Likewise, Gauss's theorem relates e to the spatial dimension and the units of e^2 are $ML^N T^{-2}$. Thus in N dimensions the dimensionless constant of Nature is proportional to

$$h^{2-N} e^{N-1} G^{\frac{1}{2}(3-N)} c^{N-4}. \tag{20}$$

It is interesting to notice that for $n = 1, 2, 3, 4$, the constants of electromagnetism, quantum theory, gravity and relativity are absent respectively. Only for $N > 4$ are they all included in a single dimensionless unit. A physical explanation of this result would be very enlightening, especially in view of the role that gauge theories with $N > 3$ may place in effecting a unification of all the fundamental forces (Salam & Strathdee 1982; Cremmer & Julia 1979; de Wit & Nicolai 1982; Cremmer *et al.* 1978; Witten 1981, 1982).

6. Mathematics

So far, we have displayed a number of special features of physics in three dimensions under the assumption that the form of the underlying differential equations do not change with dimension. One might suspect the form of the laws of physics to be special in three dimensions if only because they have been constructed solely from experience in three dimensions. If we could live in a world of seven dimensions perhaps we would end up formulating its laws in forms that made seven dimensions look special. One can test the strength of such an objection to some extent by examining whether or not 3 and (3+1) dimensions lead to special results in pure mathematics where the bias of the physical world should not enter. Remarkably, it does appear that low-dimensional groups and manifolds do have anomalous properties. Many general theorems remain unproven or are untrue only in the case of $N = 3$; a notable example is Poincaré's theorem that a smooth N dimensional manifold with homotopy type S^N is homeomorphic to S^N. This theorem is true if $N \neq 3$ and the homeomorphism can be replaced by a diffeomorphism if $N = 1, 2, 5$ or 6 (the $N = 4$ case is open). Other examples of this ilk are the problem of Schoenflies and the annulus problem; each has unusual features when $N = 3$. In addition, the low-dimensional groups possess many unexpected features because of the 'accidental' isomorphisms that arise between small groups. The twistor programme (see, for example, Penrose 1977), takes advantage of some of these features unique to (3+1)-dimensional space–times.

There is one simple geometrical property unique to three dimensions that plays an important role in physics: universes with three spatial dimensions possess a unique correspondence between rotational and translational degrees of freedom. Both are defined by only three components. In geometrical terms this dualism is reflected by the fact that the number of coordinate axes, n, is only equal to the number of planes through pairs of axes, $\frac{1}{2}n(n-1)$, when $n = 0$ or 3. These features are exploited in physics by the Maxwell field. In an $(n+1)$-dimensional space–time, electric, $\boldsymbol{E}$, and magnetic, $\boldsymbol{B}$, vectors can be derived from an $(n+1)$ dimensional potential A_i. The field $\boldsymbol{B}$ is derived from $\frac{1}{2}n(n-1)$ components of curl A_i while the $\boldsymbol{E}$ field derives from the n components of $\partial A_i/\partial t$. Alternatively we might say that in order to represent an antisymmetric second rank tensor as a vector, the $\frac{1}{2}n(n-1)$ independent components of the tensor must equal the spatial dimension, n. So the existence of axial vector representations of quantities like the magnetic vector $\boldsymbol{B}$ and the structure of electromagnetic fields is a consequence of the tri-dimensional nature of space.

There also exists an interesting property of Riemannian spaces that has physical relevance: in an $(n+1)$-dimensional manifold the number of independent components of the Weyl tensor if zero for $N \leqslant 2$, and so all the 1, 2 and 3 dimensional space–times will be conformally related and they will not contain gravitational waves. The non-trivial conformal structure for $N = 3$ leads to the properties of general relativity.

As a final example where the mathematical consequences of dimensionality spill over into areas of physics we should mention the theory of dynamical systems, or ordinary differential equations,

$$\dot{\boldsymbol{x}} = F(\boldsymbol{x}); \quad \boldsymbol{x} = (x_1, \ldots, x_N). \tag{21}$$

The solution of the system of equations (21) corresponds to a trajectory in an N-dimensional phase space. In two dimensions the qualitative behaviour of the possible trajectories is completely classified. As trajectories in two dimensions cannot cross without intersecting, the possible stable asymptotic behaviours are simple: after large times trajectories either approach a stable focus (stationary solution) or a limit cycle (periodic solution). However, when $N \geqslant 3$, trajectories can behave in a far more exotic fashion. They are now able to cross and develop complicated knotted configurations without intersecting. All the possible behaviours as $t \rightarrow \infty$ are not known for $N \geqslant 3$. When $N \geqslant 3$ it has been shown (Ruelle & Takens 1971; Plykin 1974; Newhouse *et al.* 1978) that the *generic* behaviour of trajectories is to approach a *strange attractor*. This is a compact region containing no foci or limit cycles and where all neighbouring trajectories diverge from each other exponentially in time whether followed forwards or backwards along trajectories, so there is sensitive dependence upon initial conditions. Whereas the simple attractors in the one- and two-dimensional phase spaces have dimension one (foci) or two (limit cycles) strange attractors have a *non-integral dimension*. This is manifested by the strange attractor possessing structure on all length scales. When magnified, any portion of the attractor in phase space is revealed to be just as detailed as was its large scale appearance. So far when we have mentioned 'dimension' in §§ 1–5 we have been referring to *topological* dimension but there exist other concepts of dimension that are more useful in practice. In the case of the strange attractor it has a non-integral fractal or Hausdorff dimension (Hausdorff 1918; Mandelbrot 1977; Russell *et al.* 1980). This concept of dimension gives a measure of the amount of information necessary to specify the structure of the attractor. Suppose we wish to cover a line with a minimum number of line segments of length ϵ, $n(\epsilon)$. We clearly need $n(\epsilon)$ of order ϵ^{-1}. Likewise to cover an area or a volume by elemental squares or cubes of side ϵ we require a minimum $n(\epsilon)$ of order ϵ^{-2} or ϵ^{-3} in each case. To cover a d-dimensional surface we require $n(\epsilon)$ of order ϵ^{-d}. In general, we can define a *fractal* dimension as

$$d = \lim_{\epsilon \rightarrow 0} \ln\,(n(\epsilon))/\ln\,(\epsilon^{-1}). \tag{22}$$

Strange attractors have non-integral values of $d < N$. This situation arises because strange attractors are not manifolds, but the product of a manifold and a fragmented set called a Cantor set. A simple example of a Cantor set can be generated from the unit line segment [0, 1] as follows: delete the middle third of the line segment and then the middle third from all the resulting segments and so on *ad infinitun*. In table 1 we list, L, the total length of the line segment surviving from the rth stage of the deletion process, the number of pieces, n, and the length of each piece.

After an infinite number of these operations the set that remains is a Cantor set. It clearly has zero measure because the total length of the removed pieces is the infinite geometric progression

$$\sum_{r=1}^{\infty} (\tfrac{2}{3})^r = 1. \tag{23}$$

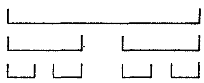
The Cantor set is nowhere dense and cannot be represented by a set of isolated points. If we choose $\epsilon = 3^{-r}$ and $n(\epsilon) = 2^r$ in (22) then we see its dimension, d, is non-integral

$$d = \lim_{r \rightarrow \infty} \ln(2^r)/\ln\,(3^r) = 0.6294\ldots. \tag{24}$$

The value of d tells us how much information is necessary to specify the location of the set to within a given accuracy.

It appears that the early evolution of the Universe, close to the initial singularity, may have possessed these exotic features and displayed chaotic structure on all length scales. In effect, general relativistic space-times can evolve as though they possess a non-integral value of d. This is the case in the Mixmaster universe (Barrow 1981 *a*, 1982; Chernoff & Barrow 1983) where an infinite number of space–time oscillations occur in any open time interval about the initial singularity at $t = 0$ and there is structure on all scales.

TABLE 1. THE CONSTRUCTION OF A CANTOR SET FROM THE INTERVAL [0, 1]

line decomposition	step	number of pieces, $n(\epsilon)$	length of each piece, ϵ	total length, L
	0	1	1	1
	1	2	$\frac{1}{3}$	$\frac{2}{3}$
	2	4	$\frac{1}{9}$	$\frac{4}{9}$
	r	2^r	3^{-r}	$(\frac{2}{3})^r$

7. Is $N > 3$?

The idea that the Universe really does possess more than three spatial dimensions has a distinguished history (Kaluza 1921; Klein 1926; Einstein & Bergmann 1938). These authors sought to associate an extra spatial dimension with the existence of electromagnetism. Under a particular symmetry assumption Einstein's equations in $(4+1)$ dimensions look like Maxwell's equations in $(3+1)$ dimensions together with an additional scalar field. Very roughly speaking one imagines uncharged particles as moving only in the $(3+1)$-dimensional subspace but charged particles move through $(4+1)$ dimensions. Their direction of motion determines the sign of their charge.

Supersymmetric gauge theories have rekindled interest in higher dimensional gauge theories that reduce to the $N = 3$ theory by a particular process of dimensional reduction. A topical example is a $(9+1)$-dimensional supergravity theory advocated by Scherk & Schwarz (1974). By analogy with the original Kaluza–Klein theories we would associate $(3+1)$ of these dimensions with our familiar space–time structure whose curvature is linked to gravitational fields and the other additional dimensions correspond to those of a set of internal symmetries. We perceive them as electromagnetic, weak and strong charges. These extra dimensions are compactified to dimensions

$$L \approx \alpha_*^{-\frac{1}{2}} L_P \tag{25}$$

where $L_P = (Gh/c^3)^{\frac{1}{2}} \approx 10^{-33}$ cm is the Planck length and $\alpha_* = 10^{-1} - 10^{-2}$ is the gauge coupling at the grand unification energy. Thus, according to such theories the Universe will be fully N dimensional (with $N > 3$) when the big bang is hotter than *ca.* 10^{17} GeV, but all except three spatial dimensions will become compactified to microscopic extent when it cools below this temperature after *ca.* 10^{-40} s. One source of interest in this scenario has been to explore whether there exist $(N+1)$-dimensional cosmological models in which such an evolution naturally occurs with three dimensions expanding while others contract. (Chodos & Detweiler 1980; Freund 1982, 1983). These authors show that one can find anisotropic cosmological solutions to $(N+1)$-dimensional general relativity in which 3 spatial dimensions expand at equal rates while the remaining $(N-3)$ spatial dimensions contract. For example, Chodos & Detweiler take a simple metric of Kasner type in $(N+1)$-dimensional space-time (see also Appelquist *et al.* 1983),

$$ds^2 = dt^2 - \sum_{i=1}^{N} a_i^2(t) dx_i^2, \tag{26}$$

then the vacuum Einstein field equation have the following solution for all N

$$a_i(t) = t^{p_i}; \quad \sum_{i=1}^{N} p_i = \sum_{i=1}^{N} p_i^2 = 1, \tag{27}$$

so at least one of the expansion rates p_i must be negative. If we require $p_1 = p_2 = p_3 \equiv p_+ > 0$ and $p_4 = p_5 = \ldots = p_N \equiv p_- < 0$ then (27) gives, in general,

$$p_+ = [\sqrt{3} + (N-3)^{\frac{1}{2}} (N-1)^{\frac{1}{2}}]/N\sqrt{3}, \tag{28}$$

$$p_- = [(N-3) - \sqrt{3}(N-3)^{\frac{1}{2}} (N-1)^{\frac{1}{2}}]/N(N-3). \tag{29}$$

For $N = 4$ we have $p_+ = -p_- = 0.5$. In the 1-2-3 dimensions this model expands at the same rate as the radiation-dominated Friedman universe. However, it is straightforward to show (Barrow, unpublished work) that models of the type (27) with monotonically increasing and decreasing axes are unstable because they possess isotropic spatial N-curvature. In general, the $(N+1)$-dimensional cosmological models possess a sequence of Mixmaster oscillations near $t \approx 0$ during which the directions of expansion and contraction are permuted in a quasi-random fashion by the anisotropic curvature (Barrow 1982). No single dimension will collapse monotonically as the Universe expands in overall volume ($(a_1 a_2 \ldots a_N) \propto t$ according to (27) for all N) and we will not in general create the situation hypothesized by Chodos, Detweiler and Freund. If dimensional compactification occurs it must have a much subtler origin.

8. The anthropic principle

Even if cosmological models like (27) were stable they would not offer a convincing explanation for the observed $(3+1)$ dimensions. After all, there is no reason why only three dimensions should be left expanding. Since other contributors (Rees, Press, Carter, this volume) are discussing some aspects of the anthropic principle (see Barrow & Tipler 1983) it is interesting to note that Whitrow (1955) first suggested an anthropic 'explanation' for why we observe space to possess three dimensions. Perhaps out of an ensemble of all possible universes of all possible dimensionalities observers can only exist in those with three space dimensions? One approach to providing circumstantial evidence in favour of the anthropic principle has been to show that life-supporting or sustaining aspects of the physical world are very sensitive to slight changes in the values of the fundamental constants of Nature (Dicke 1957, 1961; Carter 1974; Carr & Rees 1979; Barrow 1981 *b*). It is obvious from our discussion above that the consequences of the equations of physics are very sensitive to their dimension because they are differential equations, but when it comes to making small changes in the values of fundamental constants like e or G one is on much shakier ground. Although a small change in either of these quantities might so alter the rate of cosmological or stellar evolution that life could not evolve, how does one know that compensatory changes in other constants might not recreate a favourable set of solutions? Suppose, for simplicity, we treat the laws of physics as a set of n ordinary differential equations that contain a set of constant parameters λ_i which we identify with the constants of physics

$$\dot{x} = F(x; \lambda_i); \quad x \in (x_1 \ldots x_n). \tag{30}$$

The structure of the physical world is represented by the solutions of this system, say x^* for the particular realization, λ_i^* of constants that we observe. Is the solution x^* stable against small changes in the parameter set λ_i^*? This is the type of question that the Ruelle–Takens theory described in § 6 is designed to answer for generic systems of equations of the form (30). It tells us that for $n \geqslant 3$ (which will certainly be the case for our model of the laws of physics) the solution x^* will become unstable to changes in λ_i away from λ_i^* past some critical value. If the original attractor at x^* was a simple non-chaotic one with integral Hausdorff dimension then our set of laws and constants are very special in λ_i space but if the original attractor was strange then there should be many other similar sets in the λ_i parameter space. Whether these attractors have anything to do with the necessary and sufficient conditions for observers is an interesting question.

The author would like to thank Dr R. Fenn for helpful discussions.

References

Appelquist, T., Chodos, A. & Myers, E. 1983 *Physics Lett.* B **127**, 51.
Barrow, J. D. 1981*a* *Phys. Rev. Lett.* **46**, 963.
Barrow, J. D. 1981*b* *Q. Jl R. astr. Soc.* **22**, 388.
Barrow, J. D. 1982 *Physics Rep.* **85**, 1.
Barrow, J. D. 1983 *Q. Jl R. astr. Soc.* **24**, 24.
Barrow, J. D. & Tipler, F. J. 1983 *The anthropic cosmological principle.* Oxford: Oxford University Press. (In the press.)
Brouwer, L. E. J. 1911 *Math. Annln* **70**, 161.
Brouwer, L. E. J. 1913 *J. Math.* **142**, 146.
Carr, B. J. & Rees, M. J. 1979 *Nature, Lond.* **278**, 605.
Carter, B. 1974 In *Confrontation of cosmological theories with observation* (ed. M. S. Longair). Dordrecht: Reidel.
Chernoff, D. F. & Barrow, J. D. 1983 *Phys. Rev. Lett.* **50**, 134.
Chodos, A. & Detweiler, S. 1980 *Phys. Rev.* D **21**, 2167.
Courant, R. & Hilbert, D. 1962 *Methods of mathematical physics.* New York: Interscience.
Cremmer, E. & Julia, B. 1979 *Nucl. Phys.* B **159**, 141.
Cremmer, E., Julia, B. & Scherk, J. 1978 *Physics Lett.* B **76**, 409.
de Wit, B. & Nicolai, H. 1982 *Physics Lett.* B **108**, 285.
Dicke, R. 1957 *Rev. mod. Phys.* **29**, 375.
Dicke, R. 1961 *Nature, Lond.* **192**, 440.
Ehrenfest, P. 1917 *Proc. Amst. Acad.* **20**, 200.
Einstein, A. & Bergmann, P. 1938 *Ann. Math.* **39**, 683.
Freund, P. G. O. 1982 *Nucl. Phys.* B **209**, 146.
Freund, P. G. O. 1983 *Physics Lett.* B **120**, 335.
Hadamard, J. 1923 *Lectures on Cauchy's problem in linear partial differential equations.* New Haven: Yale University Press.
Handyside, J. 1929 (transl.) *Kant's inaugural dissertation and early writings on space*, pp. 11–15. Chicago: Open Court.
Hausdorff, F. 1918 *Math. Annln* **79**, 157.
Hurewicz, W. & Wallman, H. 1941 *Dimension theory.* New Jersey: Princeton University Press.
Kaluza, T. 1921 *Sber. preuss. Akad Wiss Phys. Math. Kl.* 966.
Klein, O. 1926 *Z. Phys.* **37**, 895.
Lieb, E. 1976 *Rev. mod. Phys.* **48**, 553.
Mandelbrot, B. 1977 *Fractals: form, chance and dimension.* San Francisco: Freeman.
Menger, K. 1928 *Dimensions theorie.* Leipzig: Teubner.
Neugabauer, O. 1975 *A history of ancient mathematical astronomy*, (part 2), p. 848. New York: Springer.
Newhouse, S., Ruelle, D., Takens, F. 1978 *Communs math. Phys.* **64**, 35.
Penrose, R. 1977 *Rep. math. Phys* **12**, 65.
Plykin, R. 1974 *Sb. Math.* **23**, 333.
Poincaré, H. 1917 *Dernières pensées.* Paris: Flammarion.
Ptolemy, C. 1907 *Opera* II, p. 265 (ed. J. L. Heiberg). Leipzig: Teubner.
Ruelle, D. & Takens, F. 1971 *Communs math. Phys.* **20**, 167.
Russell, D. A., Hanson, J. D. & Ott, E. 1980 *Phys. Rev. Lett.* **45**, 1175.
Salam, A. & Strathdee, J. 1982 *Ann. Phys.* **141**, 316.
Sarton, G. 1959 *History of science*, vol. 1, pp. 438–9. New York: Norton.
Scherk, J. & Schwarz, J. H. 1974 *Nucl. Phys.* B **81**, 118.
Sneddon, I. N. & Thornhill, C. K. 1949 *Proc. Camb. phil. Soc.* **45**, 318.
Weyl, H. 1922 *Space, time, matter*, p. 284. New York: Dover.
Whitrow, G. J. 1955 *Br. J. Phil. Sci.* **6**, 13.
Witten, E. 1981 *Nucl. Phys.* B **186**, 412.
Witten, E. 1982 *Nucl. Phys.* B **195**, 481.

MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES
THE ROYAL SOCIETY A
PHILOSOPHICAL TRANSACTIONS OF