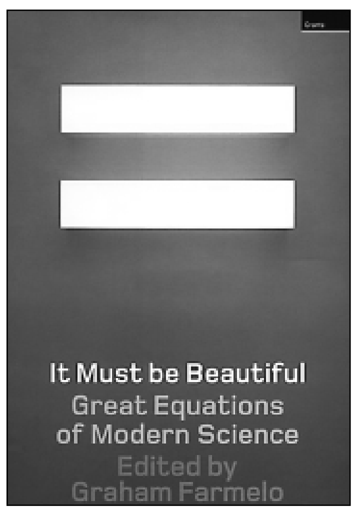
It Must Be Beautiful: Great Equations of Modern Science

Reviewed by William G. Faris

It Must Be Beautiful: Great Equations of Modern Science
Graham Farmelo, editor
Granta Publications
London, 2002
$25.00, ISBN 1-86207-479-8

It Must Be Beautiful describes eleven great equations of modern science. (It is not claimed that these are the only great equations.) To qualify as great, the equation must summarize a key principle of a domain of knowledge and have significant implications for scientific and general culture. The book is thus organized around eleven essays, each devoted to one equation. The authors of the individual essays typically state the equation (but just once, and often in an appendix), explore what it says in nontechnical terms, describe its discovery and discoverer, and explain its importance. In some cases they describe their own personal involvement. They write for the general reader; there are no technical derivations either of the equations or of their consequences.

The authors are divided among science journalists, academic writers on science, and practicing scientists. The latter category includes Roger Penrose, Frank Wilczek, John Maynard Smith, Robert May, and Steven Weinberg (who contributes a brief afterword). The writing is good; this is a book you can recommend to friends who may wonder what theoretical science is about.



This review will not attempt to comment on all the chapters. In particular, it will skip the Schrödinger equation, evolutionary game maps, and the balance equations for ozone concentration in the atmosphere. The remaining equations include five from physics, one from information theory, one from astronomy, and one from biology. The treatment in the review will be considerably more mathematical than that of the book. The plan is to describe each equation and then provide additional commentary, sometimes making connections that were not possible for authors writing independent chapters of a book.

Einstein's $E = mc^2$

The first equation is Einstein's $E = mc^2$ from his special theory of relativity. This equation relates the energy E of a particle at rest to its mass m through a conversion factor involving the speed of light c. The release of energy from atoms has obvious social consequences; part of the essay in this chapter of the

William G. Faris is professor of mathematics at the University of Arizona and is currently visiting Courant Institute of Mathematical Sciences, New York University. His email address is `faris@math.arizona.edu`.

book is devoted to this aspect of Einstein's discovery.

This way of writing the equation obscures the underlying four-dimensional geometry. The energy E is the first component of a vector $(E, c\mathbf{p}) = (E, cp_1, cp_2, cp_3)$. Here $\mathbf{p}$ is the vector that describes the momentum of the particle. The relation for a particle in motion is the equation of a hyperbola:

$$E^2 - (c\mathbf{p})^2 = (mc^2)^2. \tag{1}$$

This more general equation exhibits the mechanism of the conversion of mass into relative motion. Here is an example. A particle of mass m at rest with energy E is converted into two moving particles of masses m' with energies E' and opposite nonzero momenta $\pm\mathbf{p}'$. Conservation of energy $2E' = E$, together with relation (1) for each of the particles, leads to the inequality $2m' < m$. The particles can fly apart only if mass is lost.

The Planck-Einstein Equation $E = \hbar\omega$

The second equation is due to Planck, but its deeper significance was realized by Einstein. If anything, it is even more important for modern science. This equation is $E = \hbar\omega$, where ω is the angular frequency (frequency measured in radians per second), and $\hbar$ is a conversion constant (rationalized Planck's constant). This equation has the same flaw; it deals with only one component of a vector. The accompanying equation for the other three components was found by de Broglie; it is $\mathbf{p} = \hbar\mathbf{k}$, where $\mathbf{k}$ is the wave number (measured in radians per meter). What these equations say is that a moving particle is described by a wave. The wave associated with a particle of energy E and momentum $\mathbf{p}$ varies in time at the corresponding angular frequency ω and in space at the corresponding wave number $\mathbf{k}$. These are fundamental ideas of quantum mechanics.

Quantum mechanics is an even greater conceptual revolution than special relativity. There is an atmosphere of mystery to the subject. It is sometimes said that quantum mechanics is about measurement, but this is strange for what should be a theory of nature. However, there is no doubt that quantum mechanics gives the explanation of much of what we experience in the world, from the solidity of a table to the brilliant color of a flame.

In quantum mechanical wave equations energy and momentum are expressed by differential operators:

$$E = i\hbar\frac{\partial}{\partial t} \tag{2}$$

and

$$\mathbf{p} = -i\hbar\nabla. \tag{3}$$

Here ∇ is the gradient operator with components $\partial/\partial x_j$ for $j = 1, 2, 3$. To see how this works, apply these differential operators to a plane wave, that is, to a function of time t and space $\mathbf{x}$ of the form $\exp(-i\omega t + i\mathbf{k} \cdot \mathbf{x})$. The result is the relations $E = \hbar\omega$ and $\mathbf{p} = \hbar\mathbf{k}$. However, the differential operator formulation is more general, since the operators may be applied to functions ψ of t and $\mathbf{x}$ that are not plane waves.

Substitute equations (2) and (3) in the nonrelativistic formula $E = 1/(2m)\,\mathbf{p}^2$, and apply the result to a complex-valued function ψ. The result is the Schrödinger equation that describes the waves associated with freely moving nonrelativistic particles. Instead, substitute equations (2) and (3) in the relativistic Einstein formula (1), and again apply the result to a function ψ. The result is the Klein-Gordon equation, which describes the waves associated with freely moving relativistic particles.

Both these equations describe spin zero particles, that is, particles without intrinsic angular momentum. In nature many particles do have an intrinsic angular momentum. This angular momentum has the seemingly paradoxical property that along each axis (as determined by the imposition of a magnetic field) it has values that are integer multiples of $\hbar/2$. The simplest such case (other than spin zero) is that of spin $1/2$ particles, which can have only the two angular momentum values $\pm\hbar/2$ along each axis. The equation for describing nonrelativistic spin $1/2$ particles is the Pauli equation, a modification of the Schrödinger equation in which the function ψ has two components. The equation for describing relativistic spin $1/2$ particles is the Dirac equation, but its discovery required a new idea.

The Dirac Equation

Dirac invented his equation in 1928 to describe the motion of electrons. The equation has had a remarkable success; it seems also to describe most of the other elementary particles that make up matter, not only protons and neutrons, but even the constituent quarks.

The idea of Dirac was to rewrite the quadratic Einstein relation (1) as a linear relation. This would seem impossible. But here is the solution:

$$\gamma^0 E + c\sum_{j=1}^{3}\gamma^j p_j = mc^2 I. \tag{4}$$

What makes it work is that this is a matrix equation. There are four matrices, γ^0, γ^1, γ^2, and γ^3. They anticommute: $\gamma^j\gamma^k = -\gamma^k\gamma^j$ for $j \neq k$. Furthermore, they satisfy $(\gamma^0)^2 = I$ and $(\gamma^j)^2 = -I$ for $j \neq 0$. In these formulas I denotes the identity matrix. All this is designed so that squaring both sides of the equation gives the quadratic Einstein relation. The anticommuting property makes the cross terms cancel.

The idea of an algebra associated with a quadratic form was known before Dirac. Such algebras are known as Clifford algebras, and a standard fact about the Clifford algebra over four-dimensional space-time is that it may be represented by 4 by 4 matrices. Four components might seem to be a nuisance, but they turned out to be of great interest for Dirac. He could interpret them as describing two kinds of particles, each with two components of spin. That is, electron spin emerged naturally from this way of writing a relativistic wave equation.

Insert (2) and (3) into (4) and apply to a function ψ. The result is the Dirac equation for the wave associated with a freely moving spin $1/2$ particle:

$$(5) \qquad \gamma^0 i\hbar \frac{\partial}{\partial t}\psi - c\sum_{j=1}^{3}\gamma^j i\hbar\frac{\partial}{\partial x_j}\psi = mc^2\psi.$$

This makes sense if the solution ψ is a function of t and $\mathbf{x}$ with values that are four-component column vectors.

This is not the ultimate version of the equation, since it describes only the motion of a freely moving particle. The real problem is to describe motion in the context of an electric potential ϕ (a function of t and $\mathbf{x}$) and a magnetic vector potential $\mathbf{A}$ (a vector function of t and $\mathbf{x}$). There is a surprisingly simple recipe for doing this, which has become known as the gauge principle. The principle is to replace the energy E by $E - e\phi$ and the momentum $\mathbf{p}$ by $\mathbf{p} - e\mathbf{A}$. Here e is the charge of the particle. In the language of quantum mechanics, the gauge principle says that the quantum mechanical equations should be expressed in terms of the differential operators

$$(6) \qquad i\hbar D_t^{\phi} = i\hbar\frac{\partial}{\partial t} - e\phi$$

and

$$(7) \qquad -i\hbar\nabla^{\mathbf{A}} = -i\hbar\nabla - e\mathbf{A}.$$

This prescription works equally well for the Schrödinger equation, the Klein-Gordon equation, and the Dirac equation. Furthermore, it suggests the structure of the Maxwell equation that governs the electromagnetic potentials and fields. The Yang-Mills equation (discussed below) is a generalization of the Maxwell equation that is based on an extension of this idea.

The Dirac equation is the fundamental equation for matter, and similarly the Yang-Mills equation is the fundamental equation for force. Both equations have a geometric flavor. However, this complacent summary omits one essential feature: these are equations, not for functions, but for more complicated objects called quantum fields. There are warnings about this complication in the book. For instance, Weinberg writes:

> The more important an equation is, the more we have to be alert to changes in its significance. Nowhere have these changes been more dramatic than for the Dirac equation. Here we have seen not just a change in our view of why an equation is valid and of the conditions under which it is valid, but there has also been a radical change in our understanding of what the equation is about.

Wilczek writes the Dirac equation in the form with potentials. He employs relativistic notation, where the time and space coordinates are treated on an equal basis. However, since he uses units with $c = 1$ and $\hbar = 1$, it is possible that readers may not make the connection with the $E = mc^2$ of special relativity and the $E = \hbar\omega$ of quantum mechanics encountered in earlier chapters.

The Yang-Mills Equation

The Yang-Mills equation was invented in 1953 to describe the forces between elementary particles. It developed from an idea that Hermann Weyl published in 1929 and that was familiar to many theoretical physicists. Yang was aware of this idea, the fact that gauge invariance determines the electromagnetic interactions. However, Yang did not know its origin. As Sutton explains in her chapter: "Yang was initially unaware that these ideas were due to Weyl, and still did not realize it when both men were at the Institute for Advanced Study in Princeton and even met occasionally. Weyl had left Germany in 1933 and taken up a position at Princeton, becoming a U.S. citizen in 1939, whereas Yang joined the Institute in 1949. It seems that Weyl, who died in 1955, probably never knew of the remarkable paper that Yang wrote with Mills—the paper that demonstrates for the first time how the symmetry

Contents of *It Must Be Beautiful*

Listed below are the book's chapters and chapter authors.

"The Planck-Einstein Equation for the Energy of a Quantum", by Graham Farmelo
"$E = mc^2$", by Peter Galison
"The Einstein Equation of General Relativity", by Roger Penrose
"Schrödinger's Wave Equation", by Arthur I. Miller
"The Dirac Equation", by Frank Wilczek
"Shannon's Equations", by Igor Aleksander
"The Yang-Mills Equation", by Christine Sutton
"The Drake Equation", by Oliver Morton
"The Mathematics of Evolution", by John Maynard Smith
"The Logistic Map", by Robert May
"The Molina-Rowland Chemical Equations and the CFC Problem", by Aisling Irwin
"How Great Equations Survive", by Steven Weinberg

of gauge invariance could indeed specify the behavior of a fundamental force."

The Yang-Mills principle can be summarized in a slogan: force is curvature. Curvature is associated with surfaces. It represents the turning or rotating that takes place when some object is transported around the closed curve bounding a surface. An example of curvature is a nonzero magnetic field. Suppose $|B|$ is the total flux of magnetic field through a particular surface. The quantum mechanical phase rotates by $(e/\hbar)|B|$ as it is transported around the boundary of the surface. Here e is the electric charge, and $\hbar$ is the rationalized Planck's constant. In macroscopic contexts this is such a huge number of rotations that it is difficult to exhibit the effect directly, but the consequences on the atomic scale are quite apparent.

The technical formulation of these ideas is in a framework of differential operators inspired by the gauge principle of formulas (6) and (7). The fundamental idea is that of a connection, which is a rule for differentiating a vector field. The vector field is defined on four-dimensional space-time, with coordinates $x^0 = ct, x^1, x^2, x^3$. However, the values of the vector field belong to some N-dimensional space having to do with the internal structure of a class of elementary particles.

The gauge invariance idea may be expressed by the principle that the basis for the N-dimensional vector space may vary from point to point in space-time. Express the vector field in such a "moving basis" as $u = \sum_k u^k \mathbf{e}_k$. The connection applied to the vector field is $d^A u = \sum_k du^k \mathbf{e}_k + \sum_k u^k\, d\mathbf{e}_k$. It is determined by functions $A^j_{k\mu}$. Their purpose is to give a way $d\mathbf{e}_k = \sum_{j\mu} A^j_{k\mu} \mathbf{e}_j\, dx^\mu$ of differentiating the basis vectors. The covariant differential of the vector field is a vector-valued 1-form:

$$d^A u = \sum_{j\mu} \left(\frac{\partial u^j}{\partial x^\mu} + \sum_k A^j_{k\mu} u^k \right) \mathbf{e}_j\, dx^\mu. \tag{8}$$

In this formula the sums over j and k range over N values corresponding to basis vectors in the vector space. The sum over μ is over the four dimensions of space-time. Roughly speaking, this first covariant differential is a generalization of the idea of gradient, except that it acts on vectors rather than on scalars, and it depends on the choice of connection. In the context of Yang-Mills the connection is called the gauge potential.

The covariant derivative may also be applied to vector-valued 1-forms. The result is a vector-valued 2-form. There is an antisymmetrization in the definition of 2-forms, and so this derivative may be thought of as a generalization of the idea of curl. However, the analog of the fact that the curl of the gradient is zero no longer holds. The measure of this failure is the curvature: the linear-transformation-valued 2-form F determined by

$$d^A d^A u = F u \tag{9}$$

for arbitrary u. Due to the antisymmetry property $dx^\nu\, dx^\mu = -dx^\mu\, dx^\nu$ of differential forms, the terms involving derivatives of u cancel. The result is

$$Fu = \sum_{kj\nu\mu} \left(\frac{\partial A^j_{k\mu}}{\partial x^\nu} + \sum_m A^j_{m\nu} A^m_{k\mu} \right) u^k \mathbf{e}_j\, dx^\nu\, dx^\mu. \tag{10}$$

The second term is a matrix product, and this is responsible for the fact that F is nonlinear in A. The curvature F describes, for each point and for each small two-dimensional surface near the point, how a vector transported around the surface must necessarily twist. In the context of Yang-Mills the curvature is called the field.

The Yang-Mills equation constrains the curvature: it says that the divergence of the field F is the current J. There is an operator δ^A that plays the role of a divergence, and the equation is

$$\delta^A F = J. \tag{11}$$

This scheme is a generalization of electromagnetism. In electromagnetism the current consists of the ordinary current and charge. The curvature is a combination of electric and magnetic field. The connection is made of electric and magnetic potentials. The electromagnetic example has one special feature: the vector space has complex dimension $N = 1$, and so the matrices commute. This means that the nonlinear term is not present in the electromagnetic case. The striking feature of the general Yang-Mills equation is that the matrices do not commute; the nonlinear term is a necessary feature, imposed by the geometry. A solution is then called a nonabelian gauge field.

The equations of physics considered thus far form a closed system. Solve the Dirac equation with connection A for a function ψ. From ψ construct the current J. Integrate the Yang-Mills divergence equation with source J to get the curvature field F. Integrate the equation for F to get back to the connection A. The results should be consistent.

Einstein's General Relativity

Einstein's general theory of relativity describes gravitational force. Again there are notions of connection and curvature, but the theory is much more geometrical. The reason is that the connection differentiates vectors, but now these vectors are tangent to space-time itself.

The fundamental quantity is the metric tensor g that attaches to each space-time point a quadratic form in these space-time tangent vectors. There is a unique symmetric connection Γ with the property that

$$d^{\Gamma} g = 0. \tag{12}$$

This is used to express the connection Γ in terms of the metric g. As before, the curvature R is determined by the connection Γ by

$$d^{\Gamma} d^{\Gamma} u = Ru \tag{13}$$

for arbitrary u.

The curvature tensor R is a two-form whose values are linear transformations. Einstein wanted to find an equation that relates R to the energy-momentum tensor T. However, this tensor is a one-form with values that are vectors. It satisfies the conservation law $\delta^{\Gamma} T = 0$; that is, it has zero divergence. He struggled with this issue for some time and made a false start, proposing an incorrect equation involving the Ricci tensor, a tensor constructed from algebraic manipulations on R. Penrose explains what happened: "This equation is indeed what Einstein first suggested, but he subsequently came to realize that it is not really consistent with a certain equation, necessarily satisfied by T_{ab}, which expresses a fundamental *energy conservation* law for the matter sources. This forced him, after several years of vacillation and uncertainty, to replace the quantity R_{ab} [the Ricci tensor] on the left by the slightly different quantity $R_{ab} - \frac{1}{2} R g_{ab}$ [a form of the Einstein tensor] which, for purely mathematical reasons, rather miraculously *also* satisfies the same equation as T_{ab}!" [The italics are in the original text.]

The Einstein tensor G represents a three-dimensional curvature obtained by averaging the two-dimensional curvature $g^{-1}R$ over planes. The way G is constructed from the curvature tensor ensures that it satisfies the identity $\delta^{\Gamma} G = 0$ necessary for consistency with the conservation law. The Einstein equation relating G and T is thus simply

$$G = -\kappa T, \tag{14}$$

where κ is a physical constant. It says that the source of the three-dimensional curvature is the density of energy-momentum.

Technical note: The averaging in the definition of the Einstein tensor is seen in the explicit formula $G^a_b = -\frac{1}{4} \delta^{ast}_{bij} R^{ij}_{st}$, where $R^{ij}_{st} = g^{ik} R^j_{kst}$, and repeated indices are summed. This works out to be $G^a_b = R^{aj}_{bj} - \frac{1}{2} R^{ij}_{ij} \delta^a_b$. The corresponding quadratic form is $G_{ab} = R^j_{abj} - \frac{1}{2} R^{ij}_{ij} g_{ab}$.

Even knowing the right equation, there is still a leap from four-dimensional geometry to the fall of an apple. In his essay Penrose helps bridge this gap by giving a picture of how the geometry of the Einstein tensor is related to the "tidal effect" of gravity.

The most obvious analogy between electric field and gravitational field turns out to be misleading. The reason is that the Yang-Mills equation has a different structure from Einstein's equation. In Yang-Mills the potential A is the connection, which is related to curves. The field F is the derivative of A, so F is the curvature, which is related to surfaces. The divergence of the field F is then related to the source J. Contrast this with Einstein's equation. The potential is the metric g. The field Γ is the derivative of g, so Γ is the connection, which is related to curves. The derivative of the field Γ is a curvature R associated with surfaces. The Einstein tensor G is constructed from an average of curvatures associated with surfaces bounding a volume, and it is this tensor that is proportional to the source T. The mathematical ingredients in the Yang-Mills and Einstein theories are similar, but they are used in different ways. The relation between these two theories is a puzzle. Penrose expresses his own attitude as follows. "We know, in any case, that Einstein's theory cannot be the last word concerning the nature of space-time and gravity. For at some stage an appropriate marriage between Einstein's theory and quantum mechanics needs to come about."

Shannon's Channel Capacity Equation

Two equations are featured in this chapter. The first equation is the definition of information. If X is a random variable with probability density $p(x)$, then the corresponding information is

$$I(X) = -\int p(x) \log p(x)\, dx. \tag{15}$$

(There is also a discrete analog of this equation, where the integral is replaced by a sum.)

The second equation is Shannon's formula for channel capacity. This gives the rate at which information can be transmitted in the presence of noise. If S is the signal strength, N is the noise strength, and W is the bandwidth of the channel, then the capacity is

$$C = W \log(1 + \frac{S}{N}). \tag{16}$$

The essay stresses the role of this theory in technology. Aleksander writes: "Shannon's equations are not about nature, they are about systems that engineers have designed and developed. Shannon's contribution lies in making engineering sense of a medium through which we communicate. He shares the same niche as other great innovators such as his boyhood hero Thomas Edison (who turned out to be a distant relative, much to Shannon's delight) and Johann Gutenberg." It is only fair to say that he has also played a role as a mathematician.

The channel capacity equation is the expression of a nontrivial theorem, but a simple example can give an idea of why it takes this form. Information is to be transmitted in a time interval of length T. The input X and the output Y are vectors of real numbers of length $2WT$. (Take this as the definition of the bandwidth W, at least for the purposes of this rough computation.) The noise Z is also such a vector. It is natural to take Z to be a vector with random components. The output Y is the input X corrupted by the noise Z in the additive form

$$Y = X + Z. \tag{17}$$

It is not so obvious that it is also reasonable to model the input X as a random vector. This is justified, however, since well-encoded information looks random to a disinterested observer. Thus each of the three vectors is to consist of mean-zero Gaussian random variables. The signal vector X has total variance ST, while the noise vector Z has total variance NT. If the signal and noise are independent, then the output vector Y has total variance $ST + NT$. The information transmitted in time T is CT. This is

$$\begin{aligned} CT &= I(Y) - I(Z) \\ &= WT \log\left(\frac{\pi e(S+N)}{W}\right) - WT \log\left(\frac{\pi e N}{W}\right). \end{aligned} \tag{18}$$

The first term $I(Y)$ is the total information in Y, while the second term $I(Z)$ is the amount of this information in Y that is useless in determining X (since it is only information about the noise Z). The difference is the information in Y that is transmitted successfully from X. The result is equivalent to that given in formula (16), and thus it illustrates the Shannon formula in a special example.

The Drake Equation for Artificial Radio Sources in the Galaxy

The Drake equation is an equation for the expected number N of radio sources in the galaxy produced by intelligent civilizations. Let r be the average rate at which such sources are produced. Let L be the average length of time that a civilization persists. Then

$$N = rL. \tag{19}$$

The importance of this equation is the research program it suggests. The rate of production r is equal to the rate of production of planets in the galaxy times the conditional probability (given that the planet exists) that there will be life on a planet times the conditional probability (given that life exists) that the life will evolve to a suitable civilization. This chain of conditional probabilities can be broken down even further. The task is identified: find numerical estimates for each of these quantities. It was a valuable insight. It has also produced something of a puzzle, since many reasonable estimates give a rather large value of N. Yet we, as yet, see no such radio emissions.

The reasoning behind this equation illustrates a more general mathematical point, the importance of balance equations and of detailed balance equations. A balance equation for a stationary probability says that the rate of probability flow from all other states into a given state is equal to the rate of probability flow from that state into all other states. A detailed balance equation is a stronger and more special condition: between every pair of states the probability flows balance.

The derivation follows a standard pattern in queueing theory. The state of the galaxy is the number of civilizations n. The average rate at which civilizations are created by random fluctuations is r. The average rate of loss for a single civilization would be $1/L$, where L is the average lifetime. If there are n civilizations, then the total rate at which civilizations dissipate is n/L. If p_n is the steady-state probability of having n such civilizations, then the detailed balance equation is

$$p_{n-1}\, r = p_n \frac{n}{L}, \tag{20}$$

and since $N = rL$, this says that

$$p_n = \frac{N}{n} p_{n-1}. \tag{21}$$

The meaning of equation (20) is that the expected rate of transitions from a state of the galaxy with $n-1$ civilizations to a state with n civilizations (by having $n-1$ civilizations and then spontaneously generating a new one at expected rate r) is equal to the expected rate of transitions from a state with n civilizations to a state with $n-1$ civilizations (by having n civilizations, choosing one of the n, and having it die out at expected rate $1/L$). Then equation (21) immediately implies that the steady-state probabilities p_n have a Poisson distribution with expectation N.

It was Einstein who found a method of exploiting detailed balance in statistical mechanics in the form of fluctuation-dissipation relations. He used these relations in a particularly ingenious way to calculate the parameters describing Brownian motion and thus find the size of atoms. Special relativity, general relativity, quantum theory, Brownian motion—enough great equations for one man.

The Quadratic Map in Ecology: Chaos

It is easy to find a function f that maps an interval into itself and has one interior maximum. It can even be a second-degree polynomial in one variable. The iteration of such a function can lead to fixed points. It can also lead to periodic cycles. Of course a cycle of period n is a fixed point of the n-fold composition f^n. But there is another possibility: irregular long range behavior, that is, chaos. May describes his excitement in introducing these ideas in ecology in a phrase of Tom Stoppard: "It's the best possible time to be alive, when almost everything you know is wrong."

What did we know that was wrong? Here is how May puts it: "Situations that are effectively unpredictable—a roulette ball whose fate, the winning number, is governed by a complex concatenation of the croupier's hand, the spinning wheel and so on—were thought to arise only because the rules were many and complicated." What the quadratic map shows is that even the simplest nonlinear dynamics can produce long-range unpredictability. There is no need for complexity.

There is order in chaos. May mentions the numerical invariants of the period-doubling approach to chaos and the renormalization-group explanation of this given by Feigenbaum. The renormalization group (actually a one-parameter semigroup T^k of transformations) consists of looking at the dynamics on a longer time-scale and then rescaling to compare with the previous scale. It is a dynamical system that acts on a class of dynamical systems. In this case the dynamical systems belong to a particular space of functions f with one interior maximum. The renormalization transformation T is defined by

$$(Tf)(x) = \frac{1}{s} f(f(sx)), \tag{22}$$

where s is a scaling parameter. The Feigenbaum fixed-point equation $Tf = f$ should qualify as a great equation. Its solution f^* is a function that encodes universal properties of the approach to chaos for other functions f. There is an irony: fixed points of T describe chaotic behavior of functions f.

Conclusion

There are threads of unity linking these great equations. The equations of physics are brought together by the idea of gauge potential or connection. The formula for information is related to the formulas for entropy in statistical mechanics and to similar quantities in quantum field theory. The renormalization-group idea is now central not only to current viewpoints on statistical mechanics and dynamical systems but to quantum field theory itself.

The authors of the chapters in this volume do a remarkable job of showing how each of the great equations is situated in a broad cultural context. The equation itself is at the center. But the geometry is something like that of a black hole; the actual equation remains nearly invisible to the general reader. One of the privileges of being a mathematician is that one is allowed a glimpse inside.