Chaos, Complexity, and Entropy

A physics talk for non-physicists

Michel Baranger

Center for Theoretical Physics, Laboratory for Nuclear Science and Department of Physics
Massachusetts Institute of Technology, Cambridge, MA 02139, USA
and
New England Complex Systems Institute, Cambridge, MA 02138, USA
MIT-CTP-3112

The twenty-first century is starting with a huge bang. For the person in the street, the bang is about a technical revolution that may eventually dwarf the industrial revolution of the 18th and 19th centuries, having already produced a drastic change in the rules of economics. For the scientifically minded, one aspect of this bang is the complexity revolution, which is changing the focus of research in all scientific disciplines, for instance human biology and medicine. What role does physics, the oldest and simplest science, have to play in this? Being a theoretical physicist to the core, I want to focus on theoretical physics. Is it going to change also?

Twentieth-century theoretical physics came out of the relativistic revolution and the quantum mechanical revolution. It was all about simplicity and continuity (in spite of quantum jumps). Its principal tool was calculus. Its final expression was field theory.

Twenty-first-century theoretical physics is coming out of the chaos revolution. It will be about complexity and its principal tool will be the computer. Its final expression remains to be found. Thermodynamics, as a vital part of theoretical physics, will partake in the transformation.

CHAOS

For theoretical physicists the revolution started a few decades ago with chaos. Chaos is a purely mathematical concept; it is an undeniable mathematical fact. We know that theoretical physics is built on mathematics, and that all theoretical physicists are applied mathematicians. The first question that I want to examine, then, is: why is it that, among all the practitioners of science, applied science, engineering disciplines, and human sciences, physicists were practically the last ones to be interested in chaos and to use it in their work? There were exceptions, of course. The people who built the large particle accelerators knew about chaos; in fact they discovered much of it. There are many other exceptions. But the majority of physicists did not know about chaos, and still does not. Engineers of many sorts, meteorologists, some types of chemists, population biologists, cardiologists, economists, and even psychologists, seized upon chaos long before the physicists did. During my teaching career at MIT, twice I introduced some simple chaos in an undergraduate course for physics majors; I wanted these future physicists to be exposed to it. But in the MIT physics department, faculty members do not teach the same course for very long. After two or three years, they move you to another course to keep you from getting stale. And so, twice the person who succeeded me took chaos out of the curriculum, saying "Sorry, I don't know anything about this stuff, I can't teach it". Once again there are big exceptions, but chaos is still not part of the average American university's physics curriculum; most students get physics degrees without ever hearing about it. The most popular textbook in classical mechanics does not include chaos. Its author apologizes meekly in his introduction, saying that the subject deserves an entire book, which is a ridiculous excuse.

Why is that?

The answer is simple. Physicists did not have the time to learn chaos, because they were fascinated by something else. That something else was 20th century physics, of course! Relativity, quantum mechanics, and their myriad of consequences. I will say more: chaos was not only unfamiliar to them; it was slightly distasteful! I know, because I felt very much that way myself. What was so distasteful about it? To explain that, I need to look back a few hundred years.

Over three centuries ago Newton and Leibniz, more or less simultaneously, invented Calculus. In doing so they provided the scientific world with the most powerful new tool in its mathematical arsenal since the discovery of numbers themselves. One way to appreciate the reach of calculus is to look at it in geometrical terms. Before calculus, geometry could handle straight lines, planes, circles, spheres, and various combinations thereof. Also, with more difficulty, ellipses, parabolas, and hyperbolas; ellipsoids, paraboloids, and hyperboloids. Perhaps a few other more complicated objects. And that was all. After calculus, any curve whatsoever, any surface whatsoever, could be handled and analyzed, provided it be sufficiently smooth.

The idea of calculus is simplicity itself. It can be presented in a few lines. Let us consider a certain function, which I write

$$y = f(x) \ .$$

This notation means that, if you give me a certain numerical value for the variable x, I have in hand a mechanism (formula, computer program, experiment, or whatever) that allows me to find the numerical value of y : y is determined by x, y is a function of x, and the mathematicians write this $y = f(x)$. I can also if I wish draw a graph of this function using two axes, the horizontal axis for x and the vertical axis for y. For each value of x I use my mechanism to find the value of y, which gives me a point on the graph. I have drawn such a graph in figure 1, assuming that it was a smooth curve, and I have drawn the same curve twice. At the top, I have tried to replace the curve by fairly large segments of straight lines: it is not a very good representation. At the bottom, I have done the same thing, this time using shorter straight segments: the representation is better. If I were to use even shorter segments, the representation would be better yet. I think that you will agree that, as I go making the straight segments shorter and shorter, a procedure which mathematicians call "going to the limit", the approximation becomes more and more perfect. That is all there is to calculus; this is the complete course. Everything else in calculus just expands on this idea.

You see that it is essential that the curve be smooth, otherwise what we just did will not work. Smoothness is the key to the whole thing. There are functions, well known to

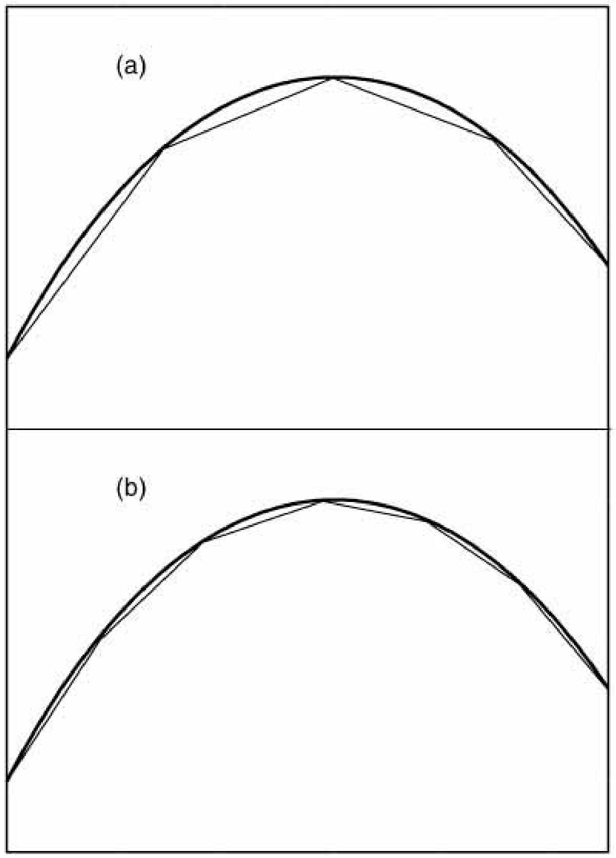

Figure 1: *The first and last calculus lesson.*

mathematicians, that are not smooth: calculus does not apply to them. But if you restrict yourself to smooth functions, you see that any one of them whatsoever can be approximated to any degree of accuracy by tiny little segments of straight lines. And since we already know everything there is to know about the geometry of straight lines, we know everything there is to know about all smooth curves! All the power of calculus follows from this.

For at least 200 years, theoretical science fed on this calculus idea. The mathematicians invented concepts like continuity and analyticity to describe smoothness more precisely. And the discovery of Calculus led to an explosion of further discoveries. The branch of mathematics so constituted, known as Analysis, is not only the richest of all the branches, but also by far the most useful for applications to quantitative science, from physics to engineering, from astronomy to hydrodynamics, from materials science to oceanography. Theoretical scientists became applied mathematicians, and applied mathematicians are people for whom analysis is second nature. Integrals, differential equations, series expansions, integral representations of special functions, etc. . . , these are the tools that calculus has provided and that are capable of solving an amazing variety of problems in all areas of quantitative knowledge.

After many decades of unbroken success with analysis, theorists became imbued with the notion that analysis was the way of the world, that all problems would eventually yield to it, given enough effort and enough computing power. This idea was not expressed explicitly; it was unconscious. But it pervades most of the science of the twentieth century. It is reflected in the choices made by university curricula and by the textbooks. The two great physical theories of the early twentieth century, Relativity and Quantum Mechanics, are both totally rooted in analysis. This idea of the invincibility of analysis was all the more powerful for being unconscious. People forgot that there were initial assumptions. The conditional truths became absolute truths. I know, because I was there myself. None of us, theoretical physicists, ever discussed the possibility that there might be practical problems where calculus did not apply. Unsmooth curves were simply "pathological phenomena of interest only to mathematicians". If you go to the bottom of this belief, you find the following. *Everything* can be reduced to little pieces of straight lines, therefore everything can be known and understood, if we analyze it on a fine enough scale. It may take a lot of work, and sometimes it may not be worth doing, but in principle we have absolute power! Yes, the enormous success of calculus is in large part responsible for the decidedly reductionist attitude of most twentieth century science, the belief in absolute control arising from detailed knowledge. Yes, the mathematicians were telling us all along that smooth curves were the exception, not the rule: we did not listen!

Chaos is the anti-calculus revolution. Chaos is the rediscovery that calculus does not have infinite power. In its widest possible meaning, chaos is the collection of those mathematical truths that have nothing to do with calculus. And this is why it is distasteful to twentieth century physicists. In terms of applications, Chaos Theory solves a wide variety of scientific and engineering problems which do not respond to calculus. This is not saying that calculus from now on will be considered outmoded, that we must focus all our attention on chaos. No, calculus retains all its power, but this power is limited. Calculus is only part of the truth. It is one member of the couple. Chaos is the bride chosen by the mathematicians to wed calculus. When calculus first saw his future bride, he really did not like her. But there will be no divorce, they are married forever, and calculus will become very fond of her, eventually. . . !

I shall outline briefly what chaos is about and how it fits with complex systems. Of all the dimensions we use to look at physical objects, the most important ones are space and time. Let us look first at chaos in space. An object which is chaotic in space is called a "fractal". There are many possible definitions of the word fractal. A very loose and general definition is this: a fractal is a geometric figure that does *not* become simpler when you analyze it into smaller and smaller parts. Which implies, of course, that it is not smooth. Simple examples of fractals have been known to mathematicians for a long time. For instance, take a straight line segment and remove its middle third. Remove the middle third of each of the two smaller segments remaining. Remove the middle third of each of the four smaller segments remaining after that. And so on ad infinitum. In the limit you get a fractal called a Cantor set. Here is another example, similar but two-dimensional. Consider the area inside an equilateral triangle. Then consider the smaller equilateral triangle obtained by joining with straight lines the centers of the three sides of the original triangle. Remove the area of this second triangle from the first. You are left with a figure consisting of three touching triangles, each of whose area is 1/4 that of the original one. Perform on each of these three triangles the same operation that was performed on the original one. And so on ad infinitum. You get a fractal called the Sierpinski triangle (see figure 2). Many people, myself included, used to think of these objects as mathematical

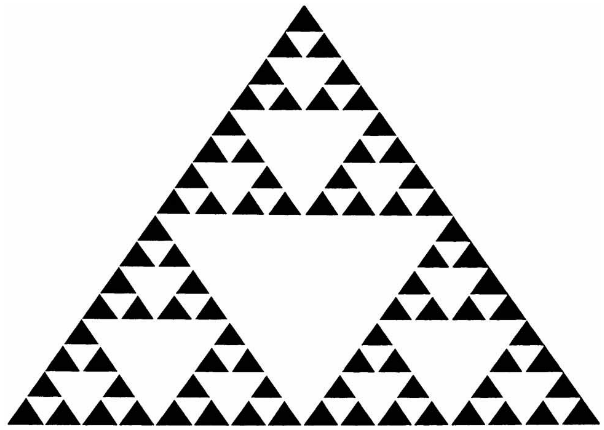

Figure 2: *The Sierpinski triangle*
(from Bar-Yam, *Dynamics of Complex Systems*).

curiosities without practical value. But we learned otherwise, especially since the 1970's. The two examples I gave are elementary, but the subject has become highly developed. There are infinite possibilities for manufacturing fractals. The famous Mandelbrot set, which is highly popular in the picture books and in the museums, is a less trivial fractal, but its definition is still amazingly simple, given the resulting complexity. It goes without saying that it is impossible to print a fractal on a piece of paper. The picture that you see on the paper is only an approximation to the real fractal. But, in the case of the Mandelbrot set, that approximation is beautiful enough, especially when you go through successive enlargements of smaller and smaller regions. For the simplest fractals like the Cantor set and the Sierpinski triangle, successive enlargements keep reproducing always the same structure: these fractals are self-similar. This is not true of the Mandelbrot set, and for it successive enlargements bring an element of surprise and delight.

Not all fractals are man-made. Nature is full of fractals. A mountain range is a fractal. A healthy old tree is a fractal. The human body is a fractal. A pattern of human settlements, a fern leaf, a pattern of earthquake faults, the sky on a partially cloudy day, the coast of Great Britain or any other section of coast, the waves on the surface of the ocean, the pattern of vegetation in the Sonoran Desert, etc., etc., all of these are fractals in the general sense given above, meaning that they do not become simpler when you examine them with an increasingly powerful microscope. Some of these things actually have a high degree of self-similarity when examined on successively finer scales: a tree, a fern, a mountain range, for instance. Others, like our bodies, do not, but they still do not simplify, as they should if they were to be treatable by calculus. We live among fractals. We feel much more comfortable among them than among the figures from our elementary geometry textbook. Isn't it strange, then, that the word fractal was not invented until 1974? There must have been some systematic opposition somewhere . . . ? And now there is also a whole series of fractals that have been designed by mathematicians or computer scientists to look like natural fractals. You often see those on your screensaver, in computer games, and in animated movies.

For a physicist like myself, there is another obvious fractal example, this one loaded with irony. We spent the whole twentieth century developing the consequences of quantum mechanics and relativity, two theories firmly based in calculus. In particular, we were eager to use these theories in the study of "elementary particles," and the result was going to be the discovery of the fundamental laws of physics and of the universe. At the beginning the elementary particles were the atoms, as proclaimed by the Greek etymology of this word. But, looking on a finer scale, we found that atoms were actually made of nuclei and electrons, and these became the new elementary particles. After we were able to refine the scale once again, we discovered that nuclei were made of protons and neutrons, plus a host of cousins and other relatives. So we changed the definition of elementary particles once more. A few decades later, thanks to accelerators of ever-increasing energy, we are able to look at the world with an even finer scale yet, and we know now that our old elementary particles are actually made of quarks of various flavors and of gluons. Will the process go on ad infinitum? We really have no good reason to say "no" any more. Particle physics, this inner sanctum of Calculus, is really one big fractal. And this fractal possesses a fair amount of self-similarity, as exemplified by the similarity between atomic spectroscopy, nuclear spectroscopy, and hadron spectroscopy.

After chaos in space, let us consider chaos in time. This is actually the more common view of chaos, and it is also where it got its name. A system whose configuration is capable of changing with time is known as a "dynamical system". A dynamical system consists of some "variables" and some "equations of motion" or "dynamical equations". The variables are any things which can vary with time. They can be multiple or single, continuous or discrete. They must be chosen in such a way that complete knowledge of all the variables determines uniquely the "state" of the system at one time. In other words, two similar systems with the same values of all the variables are in identical configurations now, and will evolve identically. The set of all possible values of the variables, i.e. the set of all possible states of the system, is called the "phase space". The present state of the system is one point in phase space. As time proceeds, this point moves in phase space. The job of the equations of motion is to determine how it moves. Given the present state of the system in phase space, the equations of motion tell you how you can calculate the state at the next instant of time. As time evolves, this point describes a "trajectory" or an "orbit" in phase space. If you know how to calculate this trajectory, you say that you have solved the equations of motion. Usually, you are given the state of the system at some initial time; this is called the "initial conditions". Then you try to calculate the trajectory which follows from these initial conditions.

The signature of time-chaos is something called "sensitivity to initial conditions". It says that, if you have two sets of initial conditions, or two points in phase space, extremely close to each other, the two ensuing trajectories, though close to each other at the beginning, will eventually diverge exponentially away from each other. Edward Lorenz, the discoverer of sensitivity to initial conditions, also called it "the butterfly effect" because, thanks to sensitivity to initial conditions, it is possible for a butterfly, by flapping its wings on some Caribbean island today, to change completely the weather pattern in Europe a month from now. If you take an arbitrary dynamical system, with equations of motion picked out of a hat, the chances are overwhelming that you will get sensitivity to initial conditions. Time-chaos in this case is the rule rather than the exception. In classical mechanics, the opposite of sensitivity to initial conditions is called "integrability". An integrable system is "multiperiodic": it can be described by a set of variables which vary periodically with time. This is true of the most common simple systems, such as the Kepler system (two heavenly bodies interacting by Newtonian gravitation) or the multidimensional harmonic oscillator. In fact it is true of all the systems discussed in most classical mechanics textbooks. These textbooks create the impression that all interesting problems are integrable, while in fact the systems that they discuss constitute a set of measure zero. And this is how generations of students have been misled!

Sensitivity to initial conditions is the death of reductionism. It says that any small uncertainty that may exist in the initial conditions will grow exponentially with time, and eventually (very soon, in most cases) it will become so large that we will lose all useful knowledge of the state of the system. Even if we know the state of the system very precisely now, we cannot predict the future trajectory forever. We can do it for a little while, but the error grows exponentially and we have to give up at some point. Does that remind you of weather prediction? Indeed: Lorenz was a meteorologist.

The connection between time-chaos and space-chaos is very close. Take a chaotic dynamical system. Pick a simple region in its phase space, such as a sphere or a cube or any other simple volume. Consider this region as a locus of possible initial conditions. Then let time flow. As each point of the region follows its trajectory, the region itself moves and changes shape. In the course of its evolution, slowly but surely the region will turn into a fractal. The fractal builds up as time progresses and becomes complete at infinite time. Every chaotic dynamical system is a fractal-manufacturing machine. Conversely, every fractal can be seen as the possible result of the prolonged action of time-chaos.

There is one aspect of chaos which took a long time to penetrate the calculus culture. It is that **chaos can be a property of very simple systems**. Until recently it was widely believed that the kind of complication which we now associate with chaos, and which of course could be observed in many places in the real world, had to be due to the interactions of large numbers of independent variables, which would make these problems very difficult to solve by any method. That was one reason for the lack of interest in these 'chaotic-looking' phenomena: they were thought to be insoluble. But this belief turned out to be entirely wrong. For a flow, i.e., a dynamical system in which the time flows continuously, three dimensions is enough to guarantee that a generic system will be chaotic in part of its phase space. For a map, which is the case where the time changes by finite steps, one single dimension is enough. A famous example of the first kind is the system studied by Lorenz: three simple equations in three variables. Lorenz wanted this to be a model of the weather, which made some people laugh, but, lo and behold, he discovered sensitivity to initial conditions. An example of the second kind is the logistic map, long used by population biologists to model seasonal variations in the abundance of a species. This simplicity feature makes chaos very different from complexity, a point on which I shall return.

This important facet of chaos, that it can be, and often is, a property of very simple systems, was a tremendous shock to me when I first found out about it. More mundanely, you can translate this as: simple questions usually have complicated answers. You ask a simple question, by giving some simple equations of motion and some simple initial conditions, and the answer, which is the trajectory in phase space, turns out to be ... chaotic! I had definitely been taught, throughout my Cartesian-imbued education in France, that simple questions must have simple answers. Anything else would be ugly, irrational, and unartistic. As you see, I have had to change.

Then there is another aspect of chaos which it shares with complexity. In both cases the dynamics have to be nonlinear. In this context, linearity means that the equations of motion do not contain any power of the variables higher than 1. In one-dimensional cases, linearity amounts to saying that the response is proportional to the stimulus. If any other power enters anywhere, the equations are said to be nonlinear. Linear equations are special: there exists a general method that solves them all exactly. They never give rise to any chaos. In a sense they are trivial; in another they are very useful, precisely because one knows exactly what to do with them. Many simple phenomena actually do follow linear equations approximately. These range from the simple pendulum, if its amplitude of oscillation is not too big, to the propagation of light in a vacuum, if the intensity of the light is not too large (which is almost always the case). Thus it is not surprising that, in the past, physicists have spent a lot of their energy on linear equations, and much less on nonlinear equations, in spite of the fact that nonlinear equations are much more general and much more prevalent in the big wide world.

And how does nonlinearity manufacture fractals and chaos? There is one and only one answer: **stretching and folding**. All flows and all maps that manufacture fractals do it by stretching and folding. Let's look at a simple example. Think of a pastry chef making a croissant. She puts down the dough and stretches it with a rolling-pin. Then she puts a layer of butter on it and folds it. She rolls and stretches it again, puts another layer of butter, and folds it again. And so on ad infinitum, or almost. What you get is an object, a delicious croissant, which is a fractal in the direction perpendicular to the table, with a very large (quasi-infinite) number of layers. This is the way all dynamical chaos works! It is easy to see how the sensitivity to initial conditions comes about. Consider two points close to each other in the initial dough. When the chef rolls the dough the first time, they get farther apart, unless they happen to be on the same vertical, which is very unlikely. Next time she rolls the dough, they get farther apart again. Eventually it is bound to happen that, in one of the folding operations, our two points end up in different layers. After that, they have completely different histories, and if they come close again it will be pure accident. All this is pretty simple and pretty obvious! But for physicists in the 1970's it was new. And it turns out to be very, very useful. And where is the nonlinearity in this? It is in the folding. Linear equations of motion have solutions whose behavior does not change with amplitude. If they start stretching, they stretch forever; they never fold. It is the nonlinearity that folds.

This is all for chaos; I must go on to complex systems. To summarize, we see that chaos destroys our reductionist dream, the dream that we have absolute power if we only know enough about the details. But chaos and calculus need each other, and they will be fruitful and multiply.

COMPLEXITY

At the present time, the notion of complex system is not precisely delineated yet. This is normal. As people work on complex systems more and more, they will gain better understanding of their defining properties. Now, however, the idea is somewhat fuzzy and it differs from author to author. But there is fairly complete agreement that the "ideal" complex systems, those which we would like most to understand, are the biological ones, and especially the systems having to do with people: our bodies, our groupings, our society, our culture. Lacking a precise definition, we can try to convey the meaning of complexity by enumerating what seem to be the most typical properties. Some of these properties are shared by many non-biological systems as well.

• 1 • *Complex systems contain many constituents interacting nonlinearly.*

Recall that nonlinearity is a necessary condition for chaos, and that almost all nonlinear systems whose phase space has three or more dimensions are chaotic in at least part of that phase space. This does not mean that all chaotic systems are complex; far from it! For one thing, chaoticity does happen with very few constituents; complexity does not. We shall return soon to the enormous difference between chaos and complexity.

• 2 • *The constituents of a complex system are interdependent.*

Here is an example of interdependence. Consider first a non-complex system with many constituents, say a gas in a container. Take away 10% of its constituents, which are its molecules. What happens? Nothing very dramatic! The pressure changes a little, or the volume, or the temperature; or all of them. But on the whole, the final gas looks and behaves much like the original gas. Now do the same experiment with a complex system. Take a human body and take away 10%: let's just cut off a leg! The result will be rather more spectacular than for the gas. I leave the scenario up to you. And yet, it's not even the head that I proposed to cut off.

• 3 • *A complex system possesses a structure spanning several scales.*

Take the example of the human body again:
Scale 1: head, trunk, limbs, . . .
Scale 2: bones, muscles, stomach, blood, nerves, . . .
Scale 3: cells, each with a nucleus, mitochondria, cytoplasm, . . .
Scale 4: chromosomes containing DNA, specialized protein molecules, each filling a special role . . .
At every scale we find a structure. This is an essential and radically new aspect of a complex system, and it leads to the fourth property . . .

• 4 • *A complex system is capable of emerging behavior.*

Emergence happens when you switch the focus of attention from one scale to the coarser scale above it. A certain behavior, observed at a certain scale, is said to be emergent if it cannot be understood when you study, separately and one by one, every constituent of this scale, each of which may also be a complex system made up of finer scales. Thus the emerging behavior is a new phenomenon special to the scale considered, and it results from global interactions between the scale's constituents. Trivial example: the human body is capable of walking. This is an emerging property of the highest scale mentioned earlier. If you study only a head, or only a trunk, or only a leg, you will never understand walking.

The combination of structure and emergence leads to **self-organization**, which is what happens when an emerging behavior has the effect of changing the structure or creating a new structure.

There is a special category of complex systems which was created especially to accommodate living beings. They are the **complex adaptive systems**. As their name indicates, they are capable of changing themselves to adapt to a changing environment. They can also change the environment to suit themselves. Among these, an even narrower category are **self-reproducing**: they know birth, growth, and death. Needless to say, we know very little that is general about such systems considered as theoretical abstractions. We know a lot about biology. But we don't know much, if anything, about other kinds of life, or life in general. There are people working on this, and it is fascinating. Let us return now to the relationship between complexity and chaos. They are not at all the same thing. When you look at an elementary mathematical fractal, it may seem to you very "complex", but this is not the same meaning of complex as when saying "complex systems". The simple fractal is chaotic, it is not complex. Another example would be the simple gas mentioned earlier: it is highly chaotic, but it is not complex in the present sense. We already saw that complexity and chaos have in common the property of nonlinearity. Since practically every nonlinear system is chaotic some of the time, this means that complexity implies the presence of chaos. But the reverse is not true. Chaos is a very big subject. There are many technical papers. Many theorems have been proved. But complexity is much, much bigger. It contains lots of ideas which have nothing to do with chaos. Chaos is basically pure mathematics, and by now it is fairly well-known. Complexity is almost totally unknown still. It is not really math. It is more like theoretical physics, or theoretical anything. Of course, once it is in good shape, it will use a lot of math, perhaps a lot of new math.

So the field of chaos is a very small subfield of the field of complexity. Perhaps the most striking difference between the two is the following. A complex system always has several scales. While chaos may reign on scale n, the coarser scale above it (scale $n-1$) may be self-organizing, which in a sense is the opposite of chaos. Therefore, let us add a fifth item to the list of the properties of complex systems:

- 5 • *Complexity involves an interplay between chaos and non-chaos.*

Many people have suggested that complexity occurs "at the edge of chaos", but no one has been able to make this totally clear. Presumably it means something like the following. Imagine that the equations of motion contain some "control" parameter which can be changed, depending on the environment (examples: temperature, concentration, intensity of some external effect like sunlight). We know that most nonlinear systems are not 100% chaotic: they are chaotic for some values of the control parameter and not chaotic for others. Then there is the edge of chaos, i.e. the precise value of the control for which the nature of the dynamics switches. It is like a critical point in phase transitions. It is the point where the long-range correlations are most important. Perhaps complex systems, such as biological systems, manage to modify their environment so as to operate as much as possible at this edge-of-chaos place, which would also be the place where self-organization is most likely to occur. It makes sense to expect self-organization to happen when there are strong long-range correlations.

Finally, there is one more property of complex systems that concerns all of us very closely, which makes it especially interesting. Actually it concerns all social systems, all collections of organisms subject to the laws of evolution. Examples could be plant populations, animal populations, other ecological groupings, our own immune system, and human groups of various sizes such as families, tribes, city-states, social or economic classes, sports teams, Silicon Valley dotcoms, and of course modern nations and supranational corporations. In order to evolve and stay alive, in order to remain complex, all of the above need to obey the following rule:

- • 6 • *Complexity involves an interplay between cooperation and competition.*

Once again this is an interplay between scales. The usual situation is that competition on scale n is nourished by cooperation on the finer scale below it (scale $n+1$). Insect colonies like ants, bees, or termites provide a spectacular demonstration of this. For a sociological example, consider the bourgeois families of the 19th century, of the kind described by Jane Austen or Honoré de Balzac. They competed with each other toward economic success and toward procuring the most desirable spouses for their young people. And they succeeded better in this if they had the unequivocal devotion of all their members, and also if all their members had a chance to take part in the decisions. Then of course there is war between nations and the underlying patriotism that supports it. Once we understand this competition-cooperation dichotomy, we are a long way from the old cliché of "the survival of the fittest", which has done so much damage to the understanding of evolution in the public's mind.

ENTROPY

Obviously the study of complex systems is going to demand that we use some kind of statistical method. We have to do mechanics and we have to do statistics at the same time. That is called **Statistical Mechanics.** But there is another name for statistical mechanics: **Thermodynamics.** Thermodynamics is about disordered energy. There is lots of that in complex systems: just think of two very popular ones, the weather and the economy. Thermodynamics is really about understanding the relationship between disordered energy and ordered energy. In physics and in chemistry, totally disordered energy is called heat. And thermo- comes from the Greek word for heat. The name thermodynamics was given by the physicists who studied these things from an empirical point of view in the middle of the 19-th century. The name statistical mechanics is due to later physicists who were studying them from a theoretical point of view, trying to derive the laws of thermodynamics from those of mechanics by using statistics.

Usually lay people have a pretty good sense of what energy is, ordered or not. So they don't have any problem grasping the first law of thermodynamics, which says that total energy, ordered plus disordered, is conserved, meaning that its numerical value does not change with time. But when it comes to the second law of thermodynamics, then people often have problems. This may be because there are many possible statements of the second law. If you choose any two of them, they often sound like they are talking about completely different things. And often they also contain some technical-sounding notions, such as the word **entropy** for instance. But entropy is simply a fancy word for "disorder". It is a quantitative measure for this extremely common concept. And actually everybody knows the second law intuitively already. It is about order compared with disorder. One easy-to-grasp statement, which is an accurate one, is the following. It concerns the time-evolution of an isolated system, which means a system lacking any kind of interaction or connection with the rest of the universe. *The spontaneous evolution of an isolated system can never lead to a decrease of its entropy (= disorder). The entropy is always increasing as long as the system evolves. If the system eventually reaches equilibrium and stops evolving, its entropy becomes constant.* In other words, an isolated system, deprived of any help from the outside, is incapable of putting its own affairs in order. At best it can let the level of disorder creep up very slowly. At worst the disorder will worsen rapidly until it becomes total, which is the equilibrium state, the state of maximum entropy. All those who have been parents of a two-year-old are totally familiar with this meaning of the second law. We see that, according to the above, for any evolving complicated system that is isolated, the entropy, which is a property of the present state of the system, never decreases with time; it may stay constant, but this happens only under ideal conditions never realized in practice; in the real world, it is always increasing. Therefore the evolution of such a system is always irreversible: if you have observed a certain evolution of your system, you know that the backwards evolution, with the same things happening in reverse order of time, can never be possible, because it would make entropy *decrease* with time. And this comes as a big shock because, in mechanics, any possible motion is also possible in the reverse direction. The second law says that in thermodynamics this is never true: for an isolated system, no thermodynamic motion is ever reversible. It is the opposite of what happens in mechanics! How could this ever come about, since we are told that thermodynamics is just mechanics plus statistics?

I shall devote the remainder of this paper to the explanation of this paradox, which is sometimes called "the paradox of the arrow of time", and to the derivation of the second law. I shall do this on the most elementary level, similar to the level of my earlier discussion of calculus. You will see that chaos is essential to the second law. If chaos does not exist, neither does the second law. The first person who realized this connection was a young Russian named Nicolai Krylov who was a graduate student in Leningrad during World War II, and who was thinking about this while on air-defense duty looking for German airplanes. He died of illness shortly afterwards at the age of 29, but his name is forever associated with the connection between chaos and thermodynamics.

Perhaps I should warn the reader that I shall be talking about a kind of chaos that is not the one that most people know about. The most important kind of chaos for the applications is called "dissipative chaos". We might call it the engineer's chaos. It is the kind of chaos which contains "strange attractors". On the other hand, the kind of chaos which I need to evoke is "conservative chaos", the physicist's chaos. It is the chaos that one encounters when working with classical mechanics, classical field theory, or any other discipline derived from Hamiltonian mechanics, with its special symplectic geometry and its many interesting conservation laws. As always, the physicists want to look at things from the most fundamental point of view, "from first principles". Conservative chaos does not have strange attractors, but it has other fractals that are equally beautiful.

Again we need to consider phase space, which I defined earlier. Thermodynamics deals only with large systems, containing many constituents. Otherwise, you would not have a distinction between ordered energy and heat. Hence phase space has a huge number of dimensions. The system may or may not be complex; that is irrelevant for the truth of the second law. If we knew the state of the system exactly, we could represent it by a point in phase space. But obviously, for a multi-constituent system, we cannot know it exactly. We have only partial knowledge, and this we can represent by a **probability distribution** in phase space. Suppose, for instance, that we know that the representative point is somewhere inside a certain volume V of phase space, but we don't know where inside this volume. Then we can represent our incomplete knowledge by a uniform probability distribution filling the volume V, with zero probability of being anywhere else (see figure 3). This is not the most general kind of probability distribution, but it will be good enough for our present purpose.

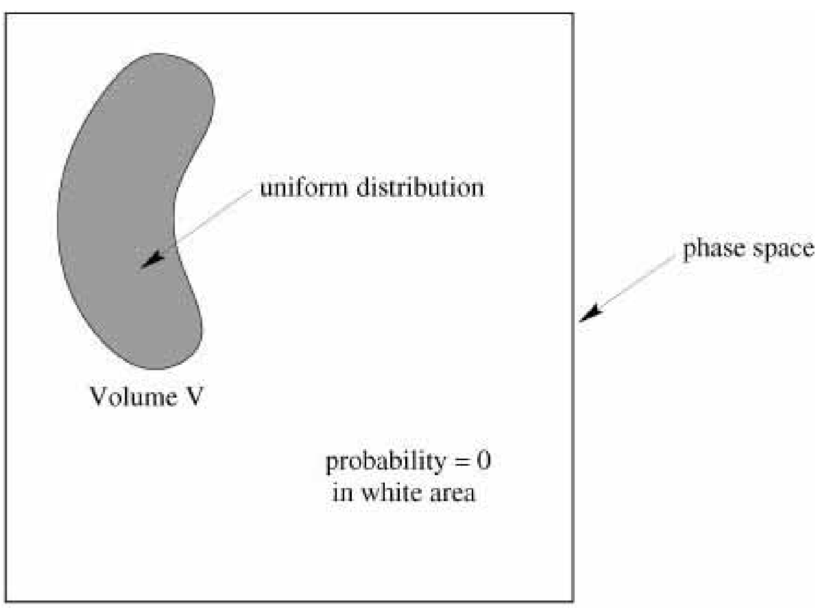

Figure 3: *An example of incomplete knowledge.*

Next, we invent a number which will represent "the amount of information" that we have about the state of the system. For instance, if the probability is uniformly distributed throughout phase space, then we know nothing and the information is $I = 0$. If we have divided phase space into two equal parts and we know which of the two halves contains the state, then the information should be, in computer language, $I = 1$ bit, because it involves making a binary choice: either it is in this half, or it is in that half. If now we divide each of the two halves into two equal parts again, and if we know which of the four quarters our state lies in, then the information should be $I = 2$ bits: we make two binary choices. And so on. The generalization of this procedure is a simple formula. If the probability is uniformly distributed throughout a volume V of phase space, and if Z is the total volume of phase space, then the amount of information (in bits) is $I = \log(Z/V)$, where the logarithm must be calculated with base 2. If the distribution is more complicated than this, there is also a more general formula, but we shall not need it now. Note that the bit is not a unit made up of mass, length, or time: information is a dimensionless quantity.

Now we define another quantity also characterizing the probability distribution, its "entropy", which we call S. The entropy is simply a measure of our lack of information. Thus we define it as follows:

$$S = \text{constant} - I \ .$$

The value of the constant is not very important, because we are more interested in the changes of entropy than in the entropy itself. In quantum mechanics the constant does have a definite value, but I won't discuss that. The unit of S is the same as the unit of I, but we have reversed the sign, so we could call it a "tib". Like I, S is dimensionless. (Note: the historical unit for S does have dimensions, but that is because thermodynamicists throw in the formula an extra dimensioned factor, Boltzmann's constant). For the case where the probability is distributed uniformly in V, our formula for I becomes the following for S

$$S = \log V + \text{another constant} \ .$$

Once again, the entropy is a measure of our lack of knowledge. If we only know that the point representing the system in phase space is inside some volume, but we don't know precisely where in this volume, then this point could be any one of a very large (actually infinite) number of points, all of which are inside the volume. Thus there is a very large number of possible states for this system and we don't know which is actually the true one. It is this large number of cognitive possibilities which is measured by the quantity we call entropy. When we say that entropy is a measure of disorder, the disorder is in our head, in our knowledge of the situation. This is actually the common meaning of the word. The qualifier "disordered", like the qualifier "random", is always used in the context of a comparison between a single, ideal, well-defined situation, and the actual situation which we consider to be only one of many different possibilities, none of which deserves to be singled out. The toys of the two-year-old should be squared away in their proper locations in drawers and boxes, but instead they are strewn about the room in places whose precise coordinates do not deserve to be noticed or recorded.

The next question, now that we have defined the entropy, is how does it vary as a function of time. The states in phase space evolve along the trajectories of classical mechanics. Each probability goes along with the state to which it belongs. Therefore the probability distribution changes with time. What happens to the entropy? The answer is very simple and it is also shocking. The entropy remains constant; it does not change at all! This is easy to see for the special case where the probability is distributed uniformly in a volume V. As time proceeds, this volume evolves, it moves around, and it changes its shape. But there is a very important theorem of classical mechanics, Liouville's theorem, which says that the numerical value of the volume never changes; it is a constant of the motion. And therefore so is the entropy, since it is essentially just $\log V$. So we now have this terrible contradiction. The second law tells us that the entropy must increase with time, except for equilibrium situations, and Liouville's theorem tells us that the entropy can never change. This paradox is the principal reason why entropy is such an interesting quantity, why the second law is such a remarkable law. Liouville's theorem, we see, is the cause of it all. If there were no Liouville's theorem, there would be no paradox !

And this is where chaos comes in. Chaos is going to resolve the paradox. True, the numerical value of the volume remains strictly constant, but according to chaos its shape changes drastically. It stretches in some directions and contracts in others. And it folds

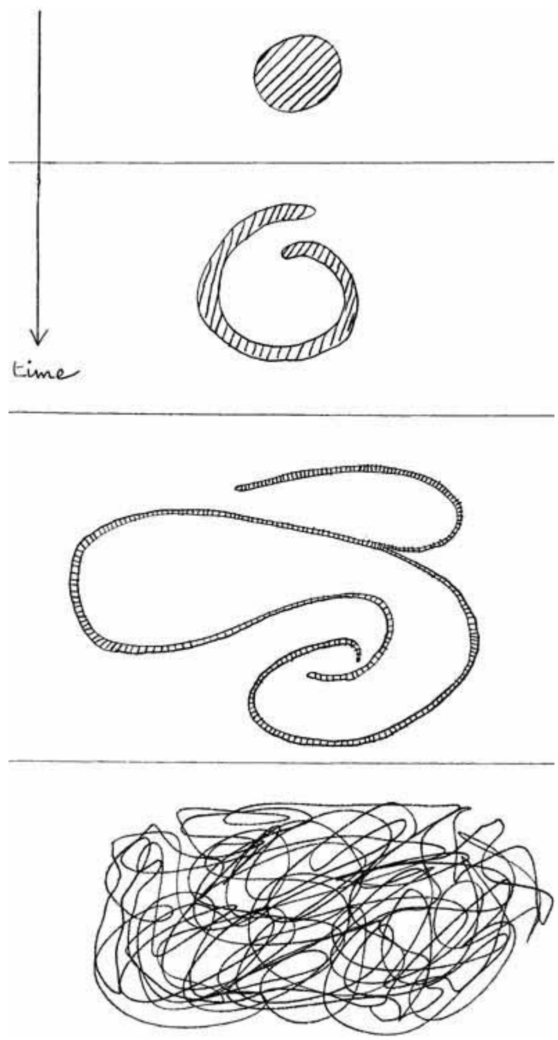

Figure 4: *The time evolution of a simple region of phase space turns it into a fractal.*

repeatedly (remember stretching and folding). Each point inside the volume tries to follow its own trajectory, but trajectories which start as neighbors diverge away eventually from each other in exponential fashion (sensitivity to initial conditions). New tendrils keep appearing and stretch everywhere. Knots of all sorts form and become more and more complicated. The structure keeps getting finer and finer. The underlying scale becomes smaller and smaller. You are witnessing a fractal in the making. In the limit of very large times, that is what you get: a fractal. And now, what about the entropy? True, the value of the volume remains the same, as guaranteed by Liouville's theorem. But it is becoming more and more difficult to see it, to follow its changes, and to measure that volume (see figure 4). Remember that, what you are looking at here, is a representation of your knowledge about the state of the system. As you watch the volume become more and more fractalized, you feel like your knowledge is slipping away from you; it cannot fit in your head any more! Eventually, as you watch your precious knowledge looking more and more like a ball of wool that the cat has gotten into, you throw up your hands in despair and you say "I can't stand this any more; I am just going to draw a nice, smooth, round bag around the whole mess" (see figure 5).

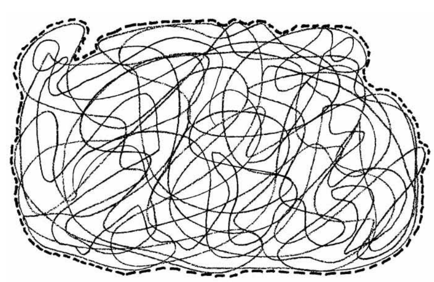

Figure 5: *Bagging the mess.*

Why do we hear all these bells and whistles suddenly? Why all the thunder and lightning? What did you just do? YOU INCREASED THE ENTROPY! Yes, you replaced the original volume by a larger volume. The original volume was full of holes (it was a fractal). You have replaced it by a smooth volume without holes, and therefore bigger. And the entropy is the logarithm of this volume: you have just satisfied the second law of thermodynamics!

Yes, you are the one who increased the entropy! It is not physics, it is not chaos, it is not Liouville: you drew the smooth volume to make your life easier, you are the one. It was chaos who manufactured the fractal, but you chose to smooth it out.

Of course, the increase in entropy does not have to be made so sudden and dramatic. It should be possible to define things so that S increases smoothly and gradually. There exists a procedure called coarse-graining which does just that. Basically coarse-graining says that, before every time at which you want to calculate the entropy, you should smooth out the details of the distribution for all scales finer that some fixed size, which should be the size beyond which you are incapable of keeping track of these details. Every such smoothing is a loss of knowledge and increases the effective volume of the distribution, hence the entropy.

The conclusion is that our dimensionless entropy, which measures our lack of knowledge, is a purely subjective quantity. It has nothing to do with the fundamental laws of particles and their interactions. It has to do with the fact that chaos messes up things; that situations that were initially simple and easy to know in detail, will become eventually so complicated, thanks to chaos, that we are forced to give up trying to know them.

One final question: if entropy is really totally subjective, why does it appear so objective to so many people? Ask any physical chemist. You will be told that entropy is a fundamental, permanent property of matter in bulk. You can measure it, you can calculate it accurately, you can look it up in tables, etc. And the answer to that new paradox is: large numbers. It is well-known that very large numbers have a way of making probabilities turn into absolute certainty. But it would take another paper to do justice to that subject.

Acknowledgments

I thank Yaneer Bar-Yam for teaching me and many others about complex systems. I am grateful to Constantino Tsallis for enlightening discussions and other help. My unforgiving editor, Marty Stock, provided much precious advice. This work was partially supported by the U.S. Department of Energy (D.O.E.) under Contract No. DE-FC02-94ER40818.

TEXTBOOKS

For the technically inclined, here are a few good books.

Steven H. Strogatz, *Nonlinear Dynamics and Chaos* (Addison-Wesley, Reading, 1994).
Undergraduate level. Mostly about dissipative chaos. Quite entertaining.

L.E. Reichl, *The Transition to Chaos*, (Springer, New York, 1992).
Graduate level. Mostly about conservative chaos. Very complete. Includes quantum chaos.

Yaneer Bar-Yam, *Dynamics of Complex Systems* (Addison-Wesley, Reading, 1997).
Invaluable. Very wide range of topics.

Roger Balian, *From Microphysics to Macrophysics*, 2 volumes (Springer, Berlin, 1991–2).
A thorough introduction to statistical mechanics.