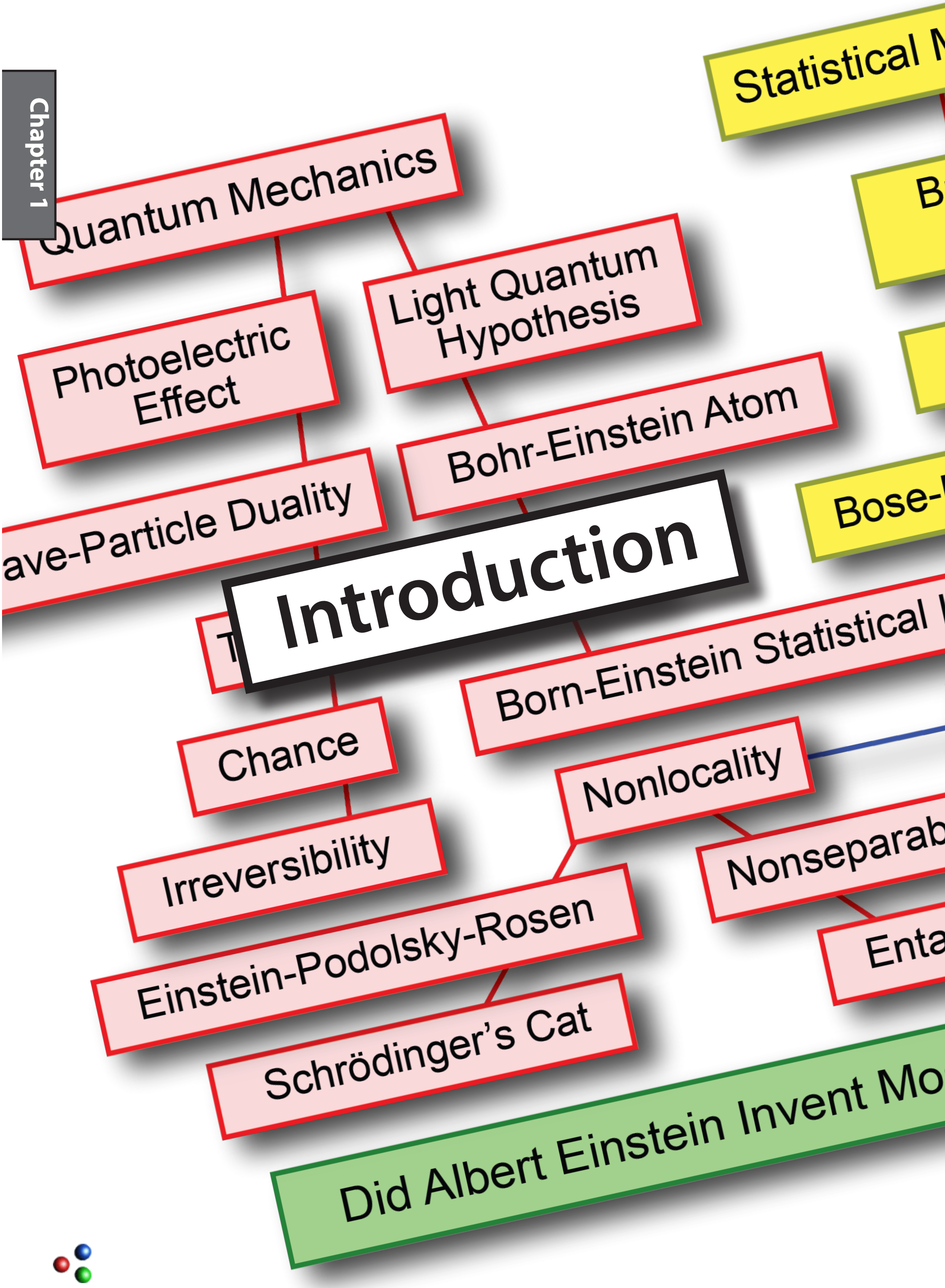


Introduction

Introduction

This book is the story of how Albert Einstein analyzed what goes on when *light* interacts with *matter* and how he discovered ontological *chance* in the process. We can show that Einstein's chance explains the metaphysical possibilities underlying the creation of all of the *information structures* in the universe.

But the story begins with a deck of cards, a pair of dice, and the multiple flips of a coin.

Around 1700, Abraham de Moivre, a French Huguenot, emigrated to England to escape religious persecution. A brilliant mathematician, he worked with Isaac Newton and other great English scientists, but he could never get an academic post, despite their excellent recommendations. To support himself, de Moivre wrote a handbook for gamblers called *The Doctrine of Chances*.

This was not the first book that calculated the odds for different hands of cards or rolls of the dice. But when de Moivre considered the flipping of a fair coin (with 50-50 odds of coming up heads and tails) he showed that as the number of flips gets large, the *discrete* distribution of outcomes approaches a *continuous* curve we call the binomial distribution, the Gaussian distribution (after the great mathematician Carl Friedrich Gauss), the "normal" distribution, or just the "bell curve," from its familar shape.

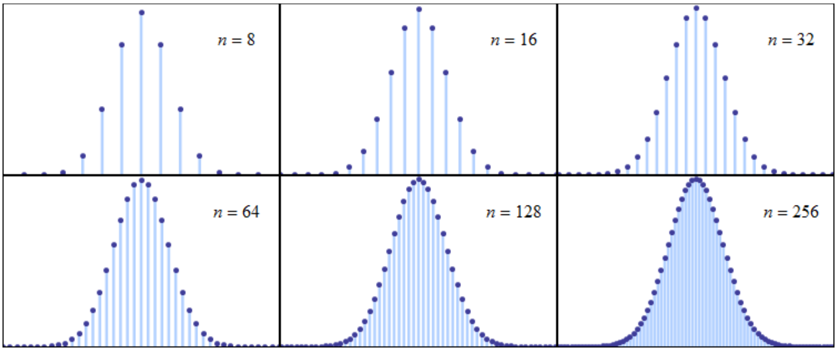


Figure 1-1. De Moivre's discovery of the continuous bell curve as a limit to a large number of discrete, discontinuous events. Each discrete event is the probability of m heads and n-m tails in n coin tosses. The height is the coefficient in the binomial expansion of $(p + q)^n$ where $p = q = ½$.

In mathematics, we can say that a finite number of discrete points approaches a continuum as we let the number approach infinity. This is the "law of large numbers" and the "central limit theorem."

But in physics, the continuous appearance of material things is only because the discrete atoms that make it up are too small to see. The analytic perfection of the Gaussian curve cannot be realized by any finite number of events.

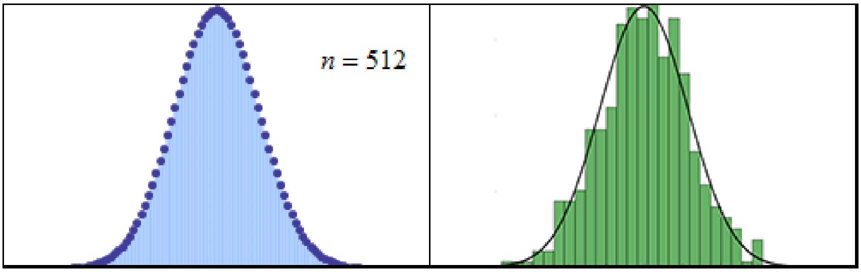


Figure 1-2. The appearance of a continuous curve and actual finite events.

Is the Nature of Reality Continuous or Discrete?

Is it possible that the physical world is made up of nothing but discrete discontinuous *particles*? Are continuous *fields* with well-defined values for matter and energy at all places and times simply theoretical constructs, averages over large numbers of particles?

Space and time themselves have well-defined values everywhere, but are these just the abstract information of the *ideal* coordinate system that allows us to keep track of the positions and motions of particles? Space and time are physical, but they are not *material*.

We use material things, rulers and clocks, to measure space and time. We use the abstract mathematics of real numbers and assume there are an *infinite number* of real points on any line segment and an infinite number of moments in any time interval. But are these continuous functions of space and time nothing but *immaterial* ideas with no material substance?

The two great physical theories at the end of the nineteenth century, Isaac Newton's classical mechanics and James Clerk Maxwell's electrodynamics, are *continuous field theories*.

Solutions of their field equations determine precisely the exact forces on any material particle, providing complete information about their past and future motions and positions. Field theories are generally regarded as *deterministic* and *certain*.

Although the dynamical laws are "free inventions of the human mind," as Einstein always said,[1] and although they ultimately depend on experimental evidence, which is always *statistical*, the field theories have been considered superior to merely statistical laws. Dynamical laws are thought to be *absolute*, based on *principles*.

We will find that the continuous, deterministic, and analytical laws of classical dynamics and electromagnetism, expressible as differential equations, are idealizations that "go beyond experience."

These continuous laws are to the discontinuous and discrete particles of matter and electricity (whose motions they describe perfectly) as the analytical normal distribution above is to the finite numbers of heads and tails. A continuum is approached in the limit of large numbers of particles, when the random fluctuations of individual events can be averaged over.

Experiments that support physical laws are always finite in number. Experimental evidence is always *statistical*. It always contains *errors* distributed randomly around the most probable result. And the distribution of those errors is often *normal*.

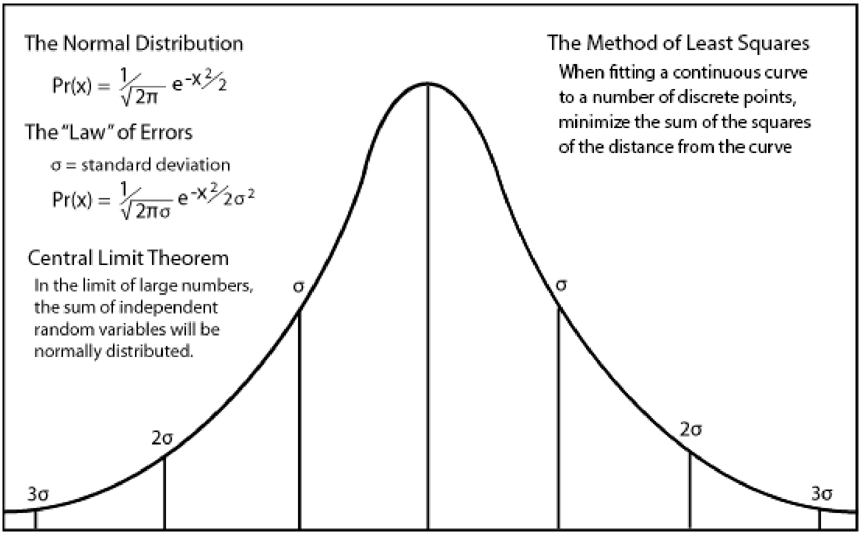


Figure 1-3. Random errors are normally distributed around the mean value.

---

[1] "Geometry and Experience," in *Ideas and Opinions*, p.234

The Absolute Principles of Physics

There are of course *absolute* principles in physics, for example the conservation laws for mass/energy, momentum, angular momentum, and electron spin. The constant velocity of light is another.

The great mathematician Emmy Noether proposed a theorem that conservation principles are the consequence of deep *symmetry* principles of nature. She said for any property of a physical system that is symmetric, there is a corresponding conservation law.

Noether's theorem allows physicists to gain powerful insights into any general theory in physics, by just analyzing the various transformations that would make the form of the laws involved *invariant*.

For example, if a physical system is symmetric under rotations, its angular momentum is conserved. If it is symmetric in space, its momentum is conserved. If it is symmetric in time, its energy is conserved. Now locally there is time symmetry, but cosmically the expansion of the universe gives us an arrow of time connected to the increase of entropy and the second law of thermodynamics.

The conservation of energy was the *first law* of thermodynamics.

The famous *second law* says entropy rises to a maximum at thermal equilibrium. It was thought by most scientists to be an absolute law, but we shall see in chapter 3 that Maxwell and Ludwig Boltzmann considered it a statistical law. Boltzmann thought it possible that a system that had reached equilibrium might spontaneously back away, if only temporarily, from the maximum. Assuming that the universe had an infinite time to reach equilibrium, he thought it might be that the non-equilibrium state we find ourselves in might be a giant *fluctuation*. Given his assumption of infinite time, even such an extremely improbable situation is at least *possible*.

In his early work on statistical mechanics, Einstein showed that small *fluctuations* in the motions of gas particles are constantly leading to departures from equilibrium. Somewhat like the departures from the smooth analytic bell curve for any finite number of events, the entropy does not rise smoothly to a maximum and then stay there indefinitely. The second law is not continuous and absolute.

The second law of thermodyamics is unique among the laws of physics because of its *irreversible* behavior. Heat flows from hot into cold places until they come to the same equilibrium temperature. The one-direction nature of *macroscopic* thermodynamics (with its gross "phenomenological" variables temperature, energy, entropy) is in fundamental conflict with the assumption that *microscopic* collisions between molecules, whether fast-moving or slow, are governed by dynamical, deterministic laws that are time-reversible. But is this correct?

The microscopic second law suggests the "arrow of time" does not apply to the time-reversible dynamical laws. At the atomic and molecular level, there appears to be no arrow of time, but we will see that Einstein's work shows particle collisions are not reversible

The first statistical "laws" grew out of examples in which there are very large numbers of entities. Large numbers make it impractical to know much about the individuals, but we can say a lot about *averages* and the probable distribution of values around the averages.

Probability, Entropy, and Information

Many scientists and philosophers of science say that the concept of entropy is confusing and difficult to understand, let alone explain. Nevertheless, with the help of our diagrams demonstrating probability as the *number of ways* things have happened or been arranged, divided by the total number of ways they might have happened or been arranged, we can offer a brief and visual picture of entropy and its important connection to information.

We begin with Ludwig Boltzmann's definition of the entropy S in terms of the number of ways W that gas particles can be distributed among the cells of "phase space," the product of ordinary coordinate space and a momentum space.

$$S = k \log W$$

Let's greatly simplify our space by imagining just two cubicle bins separated by a movable piston. Classical thermodynamics was developed studying steam engines with such pistons.

Now let's imagine that a thousand molecules are dropped *randomly* into the two bins. In this very artificial case, imagine that they all land up on the left side of the piston. Assuming the probabilities of falling into the left or right bin are equal, this is again the binomial expansion with $(p + q)^{1000}$ with $p = q = ½$. All molecules on the left would have probability $(1/2)^{1000}$. This is of course absurdly improbable if each events were random, but steam engines do this all the time, and calculating the improbability gives us a measure of the machine's available energy.

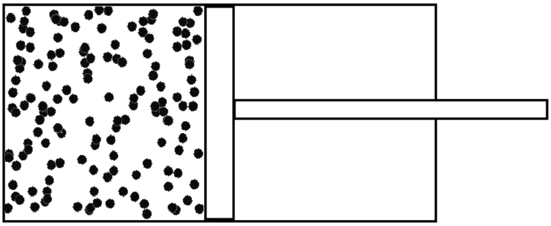
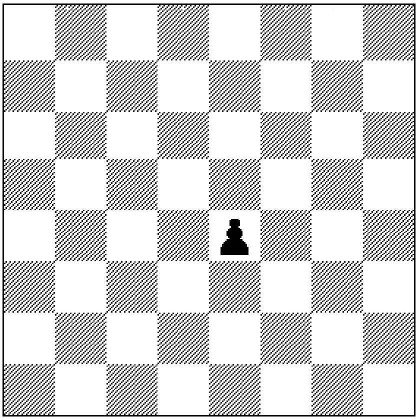


Figure 1-4. An ideal piston with gas on the left and a perfect vacuum on the right.

To see how this very improbable situation corresponds to very low entropy, how low entropy corresponds to maximum information, and how low entropy means energy available to do work, let's consider the number of yes/no questions needed to figure out the chessboard square where a single pawn is located.

1) Is it in the top half? No.
   Of the remaining half,
2) is it in the left half? No.
   Of the remaining half,
3) Is it in the right half? No.
   Of the remaining half,
4) Is it in the top half? Yes.
   Of the remaining half,
5) Is it in the left half? Yes.
   Of the remaining half,
6) Is it in the top half? Yes.



In Claude Shannon's 1948 theory of the communication of information, the answer to a yes/no question communicates one bit (a binary digit can be 1 or 0) of information. So, as we see, it takes 6 bits of information to communicate the particular location of the pawn on one of the 64 possible squares on the chessboard.

Shannon and his mentor, the great mathematical physicist John von Neumann, noticed that the information I is the logarithm of the number of possible ways W to position the pawn. Two raised to the 6th power is 64 and the base 2 logarithm of 64 is 6. Thus

$I = log_2 W$ and $6 = log_2 64$

The parallel with Boltzmann's entropy formula is obvious. His formula needs a constant with the physical dimensions of energy divided by temperature (ergs/degree). But Shannon's information has no physical content and does not need Boltzmann's constant k. Information is just a dimensionless number.

For Shannon, entropy is the *number of messages* that can be sent through a communications channel in the presence of noise. For Boltzmann, entropy was proportional to the *number of ways* individual gas particles can be distributed between cells in phase space, assuming that all cells are equally probable.

So let's see the similarity in the case of our piston. How many ways can all the 1000 gas particles be found *randomly* on the left side of the piston, compared to all the other ways, for example only 999 on the left, 1 on the right, 998 on the left, 2 on the right, etc.

Out of 2^{1000} ways of distributing them between two bins, there is only *one way* all the particles can be on the left. [2] The logarithm of 1 is zero ($2^0 = 1$). This is the minimum possible entropy and the maximum of available energy to do work pushing on the piston.

Boltzmann calculated the likelihood of random collisions resulting in the *unmixing* of gases, so that noticeably fewer are in the left half of a 1/10 liter container, as of the order of $10^{10^{10}}$ years. [3] Our universe is only of the order of 10^{10} years old.

The most probable cases distribute gas molecules nearly equally on each side of the piston. A hole in the piston that leaked gas would reduce the available energy to push the piston and do work. Low

---

[2] 1000! (factorial) is 1000 x 999 x 998 ... x 2 x 1. (really big)

[3] *Lectures on Gas Theory, p.444* Appendix F.

entropy is a measure of available energy. Maximum entropy means no available energy - and no information.

Low entropy is highly improbable, unless there is a *providential* creative process at work in the universe. There are a few such processes. Information philosophy has identified them.[4]

Understanding and Visualizing *Phase Space*

To visualize how matter and radiation are distributed in the universe, and ultimately in the material of our own world and the matter in our own bodies, we can start by thinking of space as made up of small cells of volume h^3, where h is Planck's constant.[5]

Now this may not be a property of space itself, but a property of how tightly we can pack material particles[6] into space. Planck's constant h is a measure of the uncertainty (or indeterminacy) in the product of position uncertainty Δx and momentum uncertainty Δp. We cannot pack particles into smaller volumes than $h^3 = \Delta x^3 \cdot \Delta p^3$.

We can pack more particles in the same easily visualizable Δx^3 position space only by adding particles with higher and higher momenta, or equivalently higher and higher energies. But where all position space cells have the same *a priori* probability, the probability of higher energy cells is diminished by $e^{-E/kT}$ or $e^{-p^2/2mkT}$.

The fact that higher momentum cells have lower "statistical weights" and thus lower populations or "occupancy numbers" was implicit in Maxwell's distribution of gas particle velocities, but $e^{-E/kT}$ is known as the "Boltzmann factor," as we shall see in chapter 4.

We are now in a position to look at the birth and evolution of the universe by asking how exactly every particle has come to be in a specific cell of phase space, controlled by what forces and whether it is there by chance or by some arrangement process.

Our Big Bang model of the universe begins with an energy density higher than physics can currently describe. From *t=0* the universe expands - as Einstein's general relativity predicts - producing more cells of position space into which particles of matter can be arranged. The standard model of particle physics predicts that quarks are pulled together by the strong nuclear force of gluons

---

[4] *Great Problems in Philosophy and Physics, Solved?* Appendix F.

[5] In chapter 4 we will see how Planck discovered *h*.

[6] There is no limit to the number of *light* particles, as Einstein found in 1925.

to form hadrons, baryons and mesons, and ultimately the familiar nuclear particles protons and neutrons. These form mostly hydrogen and some helium nuclei. After some 380,000 years, the matter and radiation have cooled down to about 5000K (the temperature of the solar surface and the color temperature of light falling on the earth).

Up to this time, the universe had been approximately in thermodynamic equilibrium, near maximum entropy. The material gas and the radiation particles were all distibuted randomly and had the same rapidly declining temperature. But now the protons and helium nuclei combined with electrons to form the first atoms.

Where free electrons had been scattering photons, maintaining the same temperatures for matter and light, these two now decouple and the radition temperature starts to fall faster then the temperature of matter. This decoupling marks the emergence of *structural information* in the universe. Had there been an intelligent observer to see, the local material gas would now appear transparent and the distant uniform cosmic background would slowly darken over the next few hundred million years to a dark sky with no visible light.

But in that sky of a "dark age" would appear spectacular new sources of light and streams of cosmic ray particles of enormous energy. The relatively weak force of universal gravitation, nearly 40 orders of magnitude smaller than the forces holding the atoms together, attracted small fluctuations in the thin material gas. As the particles condense, they heat up so they can start to again occupy the higher energies in momentum space.

The crush of gravitation is so great, internal temperatures so high, that the primordial nuclei combine to form all the heavier elements of today's cosmic composition. We can call these events stars, but they are mostly short lived and explode, dispersing their new heavy elements. The biggest left behind cores that became quasars or even small black holes.

These first *visible* information structures we can call *gravitational structures*. For some, radiation pressure balanced the gravitational force and they became long-lived *stars*. Very small gravitational condensations may never have had thermocuclear interiors and exist today as *planets* and smaller objects. In these, chemical forces balanced the gravitation.

For the planets, stars, and galaxies, we can say that all the material particles initially arrive at their particular positions in phase space as a result of random chance. While stars are gaseous objects, in planets there are great numbers of similar particles that aggregate into the solid state and move together. We can think of them as chemical or *geological* structures.

On the surface of planets orbiting stars in a narrow range of distance that allows liquids like water to form, we find the development of *biological* information structures. In these, and in their artifacts, the positions and momenta of particles in phase space are under the control of processes managed by other information structures.

Biological structures are full of information and are quite low entropy. As Erwin Schrödinger noted in his famous essay, "What Is Life," humans, animals, and the plants they ultimately depend on, all "feed on negative entropy" from sunlight.We can now see that the low entropy of sunlight is ultimately the product of gravitational collapse, which, in turn, was made possible because Einstein's expansion of the universe increased the amount of coordinate space.

The universe expansion produced many more phase space cells into which particles can be arranged, initially randomly by gravitation and chemical forces, but later in living things under the control of the information structures and processes they contain. The particular atoms in a living macromolecule like RNA or DNA are selected by chance, but the type of atom at each location is controlled by quantal processes of incredible speed and accuracy.

Similarly, the next amino acid assembled into a growing polypeptide chain is the first correct one that bumps by chance into the ribosome and is recognized by the molecular conformation of its transfer RNA. There are 60-plus tRNA codon sequences *randomly* bumping the ribosome at hundreds of times per second and traveling at hundreds of mile per hour. Despite the random chance adding new components, the quantal assembly of atomic and molecular structures proceeds with phenomenal accuracy.

We must change the impression that quantum events can only contribute noise and not new information structures.

The Cosmic Creation Process

We can thus summarize the creation and evolution of information structures in four important cosmic epochs:

1. From the origin to the formation of atoms (~380,000 years).
2. The formation of galaxies, stars, and planets (~400 million years).
3. The creation and evolution of life on Earth (~9 billion years).
4. The evolution of the human mind.

Information philosophy (actually information physics and biology) has identified the two steps in the process needed to create any new information structure.

1. A Quantum Step.

Whenever matter is rearranged to create a new information structure, the quantum binding forces involve a collapse of the wave function that introduces an element of *chance*. Without *alternative possibilities*, no new information is possible, and there would be just one possible future. With alternative possibilities, things could have been otherwise, and the future is open, allowing our choices to help create the universe.

2) A Thermodynamic Step.

A genuinely new information structure reduces the local entropy. It cannot be stable unless it transfers away enough positive entropy to satisfy the second law of thermodynamics, which says that the total entropy (disorder) must always increase.

With the emergence of self-aware organisms and the creation of extra-biological information stored in the environment, the same two-step information-generating process underlies human communication, consciousness, free will, and creativity.

Einstein not only saw clearly the possibility of "free creations of the human mind," he also contributed many that we hope this book will remind us of. Historians of science correct his record. And interpreters of quantum mechanics may now give us a comprehensible picture to replace an incomprehensible one.