UNITED NATIONS EDUCATIONAL, SCIENTIFIC AND CULTURAL ORGANIZATION
INTERNATIONAL ATOMIC ENERGY AGENCY
INTERNATIONAL CENTRE FOR THEORETICAL PHYSICS
I.C.T.P., P.O. BOX 586, 34100 TRIESTE, ITALY, CABLE: CENTRATOM TRIESTE

H4.SMR/1011 - 12

Fourth Workshop on Non-Linear Dynamics and Earthquake Prediction

6 - 24 October 1997

Determinism, Chaos, Randomness

H. MORITZ

Graz Technical University
Institute of Physical Geodesy
Graz, AUSTRIA

MAIN BUILDING STRADA COSTIERA, 11 TEL. 2240111 TELEFAX 224163 TELEX 460392 ADRIATICO GUEST HOUSE VIA GRIGNANO, 9 TEL. 224241 TELEFAX 224531 TELEX 460449
MICROPROCESSOR LAB. VIA BEIRUT, 31 TEL. 2249911 TELEFAX 224600 TELEX 460392 GALILEO GUEST HOUSE VIA BEIRUT, 7 TEL. 2240311 TELEFAX 2240310 TELEX 460392
ENRICO FERMI BUILDING VIA BEIRUT, 6 (TELEPHONE, FAX AND TELEX THROUGH MAIN BUILDING)

Determinism, Chaos, Randomness

Helmut Moritz
Graz University of Technology

Abstract. This paper reviews two separate but related problems, which both go back to Henri Poincaré: the "method of arbitrary functions", which provides a reduction of physical probabilities to instability and symmetry, and the general theory of nonlinear dynamic systems, popularly known as Chaos Theory.

1 Introduction

Let us start with three quotations:

> An intelligent being which, for some given moment of time, knew all the forces by which nature is driven, and the relative position of the objects by which it is composed (provided the being's intelligence were so vast as to be able to analyze all the data), would be able to comprise, in a single formula, the movements of the largest bodies in the universe and those of the lightest atom: nothing would be uncertain to it, and both the future and the past would be present to its eyes. The human mind offers in the perfection which it has been able to give to astronomy, a feeble inkling of such an intelligence.
>
> Pierre Simon de Laplace (1799)

> We collectively wish to apologize for having mislead the general educated public by spreading ideas about the determinism of systems satisfying Newton's laws of motion that, after 1960, were proved incorrect.
>
> Sir James Lighthill (1994)

> Imagine the figure formed by these two curves and their infinitely many intersections ...; these intersections form a kind of meshwork, tissue, or infinitely dense network ... One is struck by the complexity of this figure which I do not even attempt to draw. Nothing is better suited to give us an idea of the complexity of the three-body problem and in general of all the problems of dynamics in which there is no uniform integral [of the motion] ...
>
> Henri Poincaré (1899)

These three quotations, each about a century apart, show:

- classical determinism (Laplace)
- deterministic chaos (Lighthill)
- the essence of chaoticity (Poincaré)

Classical determinism, based on Newtonian mechanics, considers the universe completely determined by stable deterministic laws wich, in principle, can be known, as well as their stable solutions, by a superhuman intelligent being ("*Laplace's demon*").

A very famous example is the *two-body problem* of classical mechanics, such as the motion of the Earth around the Sun, along a Kepler ellipse, which is perfectly stable from time $= -\infty$ to time $= +\infty$.

Surprisingly enough, Poincaré showed that even a relatively small perturbing body such as the Moon (*three-body problem*) changes the classical simplicity in a completely unexpected and dramatic ("chaotic") fashion. The trajectories, far from forming regular geometric curves, show an irregular meshwork which Poincaré "does not even attempt to draw". Only present-day computers can plot such an intricate mesh (Figure 1). It can be shown that, in a certain sense, chaotic behavior is the rule and the simple

Figure 1: Poincaré's "chaotic" behavior of trajectories (after Herrmann 1994)

classical motions, which Laplace had in mind, are rather exceptions. This is the reason for Lighthill's impressive statement quoted above.

2 Instability

Stability is a continuous dependence of the trajectories from the *initial conditions*: small causes lead to small effects. Instability means that small causes may produce large effects.

Again nobody can express the situation better than Poincaré (1908). From the English translation (Appendix B) of his book we quote from p. 67:

> "We will select unstable equilibrium as our first example. If a cone is balanced on its point, we know very well that it will fall, but we do not know to which side; it seems that chance alone will decide. If the cone were perfectly symmetrical, if its axis were perfectly vertical, if it were subject to no other force but gravity, it would not fall at all. But the slightest defect of symmetry will make it lean slightly to one side or other, and as soon as it leans, be it ever so little, it will fall altogether to that side. Even if the symmetry is perfect, a very slight trepidation, or a breath of air, may make it incline a few seconds of arc, and that will be enough to determine its fall and even the direction of its fall, which will be that of the original inclination.
>
> A very small cause which escapes our notice determines a considerable effect that we cannot fail to see, and then we say that that effect is due to chance. If we knew exactly the laws of nature and the situation of the universe at the initial moment, we could predict exactly the situation of that same universe at a succeeding moment. But, even if it were the case that the natural laws had no longer any secret for us, we could still only know the initial situation *approximately*. If that enabled us to predict the succeeding situation *with the same approximation*, that is all we require, and we should say that the phenomenon had been predicted, that is governed by laws. But it is not always so; it may happen that small differences in the initial conditions produce very great ones in the final phenomena. A small error in the former will produce an enormous error in the latter. Prediction becomes impossible, and we have the fortuitous phenomenon."

Another example of instability in classical mechanics is the fall of raindrops, considered small rigid spheres (Figure 2). This is of course unrealistic, but it is simple, and realistic liquid raindrops would lead to essentially the same result. Again, neighboring trajectories suffer an instability at the top of the roof: a raindrop may go down the left-hand side of the roof, whereas an arbitrarily close second raindrop may take the right-hand side of the roof.

Still another example is the throw of a coin: on the first throw, the coin may show "head", whereas on another throw (even if we try to reproduce as well as possible the "initial conditions" provided by the throwing hand) may show "tail".

On the average, both sides of the coin are equally likely: performing a great number of coin tosses, head and tail will have equal probability $p = 1/2$, *provided the coin is perfectly symmetric.*

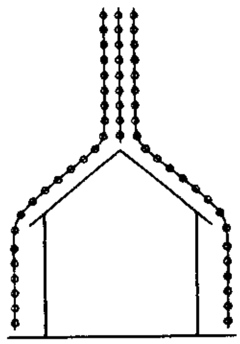

Figure 2: Trajectories of raindrops falling on a roof

Quite similar is the throw of a die: on repeatedly throwing a perfectly symmetric die, each face will have equal probability.

Thus we see that the initial conditions and the laws of mechanics become almost irrelevant for the statistical outcome, and *symmetry takes over*; cf. (Moritz 1995, p. 84); Appendix A.

3 The Implication of Roulette

In the beginning ... there was Poincaré.

E. Atlee Jackson

The game of roulette is well-known to gamblers in gambling casinos. Since scientists are not usually gamblers, let me briefly sketch the very simple principle: The roulette wheel is divided into a large number of equal alternate red and black sectors (Fig. 3). It rotates with little friction, until it comes finally to a stop. This stop evidently is again practically independent of the initial condition (the *tourneur* causing the wheel to spin, whereupon the wheel continues to rotate freely). The wheel stops and a fixed needle points either to a red or a black section, determining whether we lose or win. With a fair wheel, red and black are evidently equally probable ($p_1 = p_2 = 1/2$).

Poincaré investigated this case (Poincaré 1908, p. 76) and similar cases mathematically in (Poincaré 1912, chapter XVI). Thus he created a general mathematical theory, known by the name of "*method of arbitrary functions*". This method was further developed by Hopf (1937) and others; an excellent review with references is (Engel 1992).

Much better known is Poincaré's role as a founder of nonlinear dynamics ("deterministic chaos") already mentioned above. It should be mentioned, however, that, together with Birkhoff (1927), Hopf (1937) was also a forerunner of chaos theory, which exploded with the work of A.N. Kolmogorov (also the founder of modern probability theory!) and his associate V.I. Arnold after 1954 and, independently, of E.N. Lorenz after 1963. Is it surprising that Poincaré (1908, p. 68; see Appendix B) already fully understood and

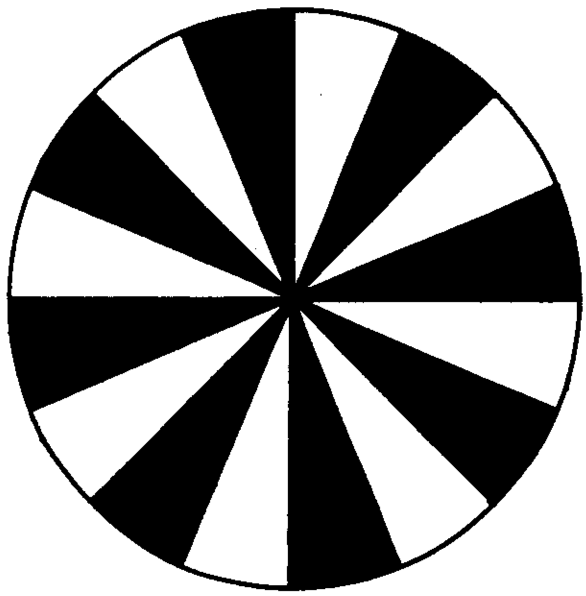

Figure 3: A roulette wheel

described meteorological instability, which was the starting point of Lorenz' pioneering investigations?

My favorite books on chaos theory are (Arnold and Avez 1968; a classic in content and readability), (Abraham and Shaw, 1992; unsurpassed simple presentation of the complex geometry without formulas!), (Hilborn 1994; excellent introduction to concepts and methods), and (Jackson 1990; comprehensive geometric treatment on high but accessible level). Herrmann (1994) and Korsch and Jodl (1994) give beautiful and highly instructive computer programs with diskettes. The game of roulette has important implications indeed!

4 The Drunkard and the Gaussian Distribution

The most important probability distribution is the *normal distribution*

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{x^2}{2\sigma^2}} \quad , \tag{1}$$

where π and e are the standard mathematical constants and σ (the dispersion) is a constant characterizing the distribution.

Its importance consists in the *Central Limit Theorem*: a random phenomenon, caused by a combination of many small random influences will be normally distributed even if the small original influences are arbitrarily distributed.

Now the roulette (of radius R) has a "limiting distribution"

$$f(x) = \frac{1}{2\pi R} = \text{const.} \quad , \tag{2}$$

that is, equal distribution on the circle, because of the very symmetry of the wheel. For the infinite straight line, however, a uniform distribution is impossible because the constant would have to be identically zero (for $R \to \infty$), in contradiction to the well-known fact that the area under the curve $f(x)$ must be 1 (corresponding to probability 1 or certainty).

Thus translational symmetry for $f(x)$ is impossible: one point ($x = 0$) must be fixed. Then which symmetry is exhibited by the Gaussian distribution (1)?

It is rotationally symmetric: rotating the curve (1) around the y-axis produces a distribution which again is a two-dimensional normal distribution. A rotation is a special case of a linear transformation. This rotational invariance and invariance with respect to dimension 1 or 2 is a particular case of two basic invariances:

1. invariance with respect to dimension (there are normal distributions in n dimensions) and

2. invariance with respect to linear transformations.

Linear or linearized laws are still of basic importance in physics. Even quantum theory is strictly linear. Einstein's General Theory of Relativity is approximately linear, although nonlinearities (e.g. "black holes") are of increasing importance. Of course, chaos theory is basically nonlinear, and so is the "fractal geometry of nature" according to B. Mandelbrot. Still, most of contemporary physics is linear or linearizable.

The linearity of (1) with respect to linear transformation is also at the root of the Central Limit Theorem: only a linearly invariant law can be a limiting distribution of a sum of many small random influences, since a sum is linear by its very definition! (This heuristic argument, however, would be a poor "proof" of the Central Limit Theorem!)

A beautiful physical interpretation of the Central Limit Theorem is Galton's Board. This apparatus is described in many books on statistics and intuitive mathematics (even the great Kolmogorov does not distain it (Kolmogorov 1988, p. 185)). Fig. 4 shows two versions found in the literature. Falling small balls are "randomly" deviated by symmetrically arranged nails (or similar obstacles). At least in the left figure, the ball performs a *random walk*, that is, every step can be to the right or to the left with equal probability. (A random walk is performed by a "perfect drunkard" who is so drunk that he cannot distinguish right or left, but still not so drunk as to be unable to walk). The similarity with the instability of Fig. 2 is evident. Poincaré is reported to have said, partly in jest no doubt, that there must be something mysterious about the normal distribution since mathematicians think it is a law of nature whereas physicists are convinced that it is a mathematical theorem ...

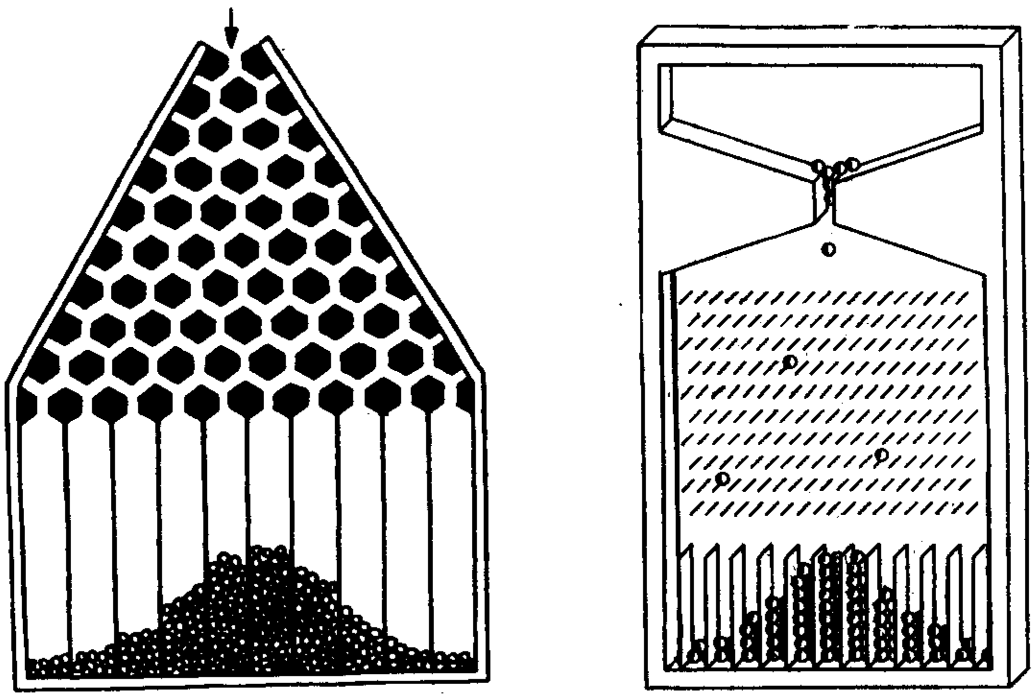

Figure 4: Two versions of Galton's Board

5 Final Remark

What is randomness? Paraphrasing a famous statement of St. Augustine about time, we may say: "If nobody asks me, I know what randomness is; if I want to explain it to somebody, then I don't know what to say."

At any rate, randomness has many aspects. Chaos theory shows that determinism is compatible with what we would call random behavior. Conversely, statistical mechanics (which we have not touched yet) derives the "deterministic" laws of thermodynamics from the "probabilistic" statistical mechanics.

For a somewhat provocative interpretation of the toss of a coin assume that we have a law which associates to any given digit 1 the occurrence of "head", and to any given digit 0 the occurrence of "tail".

Take any real number in the interval (0, 1) and represent it as a *binary number* (known to everyone in our computer age):

$$x = 0.011000100111110100\ldots \tag{3}$$

If we prescribe such a real number *exactly* by means of an infinite number of binary digits, the sequence of coin tosses is perfectly, "deterministically", known in advance ...

The basic problem in this example is of course the impossibility of determining or fixing *exactly* the boundary condition (already pointed out by Poincaré above), in our case the number (3). This, however, is a general feature of all physical measurements:

they can never be "mathematically exact" (whatever this means). This is the reason for the probability of coin tossing, dice throwing, roulette, etc.

By the way, the normal distribution is *the* standard tool of the theory of measuring errors and has been developed by Laplace and Gauss for this purpose.

As a matter of fact, quantum theory is essentially probabilistic (cf. Ruhla 1992, chapter 7). Are at least some of the small "random influences" on measuring errors due to random quantum fluctuations (Moritz 1995, pp. 254–255)?

Finally we mention that probabilistic methods are very successfully applied in the most austere and "pure" branch of mathematics: *number theory*. As the beautiful booklet (Kac 1959, p. 53) expresses it, "primes play a game of chance". Normal distribution even plays a role in number theory. One of the basic concepts of statistical mechanics and chaos theory, *ergodicity* (Birkhoff 1927; Hopf 1937; Arnold and Avez 1968), occurs with continued fractions (Kac 1959, p. 89).

And very delicate number-theoretic property decide about stability or instability of trajectories of nonlinear dynamic systems ("KAM theorem", after A.N. Kolmogorov, V.I. Arnold and J. Moser) ...

We are back at the beginning: it is time to stop.

References

Arnold V I and Avez A (1968) Ergodic Problems of Classical Mechanics, Benjamin, New York.

Abraham R H and Shaw C D (1992) Dynamics: The Geometry of Behavior, 2nd ed, Addison-Wesley, Redwood City, Cal.

Birkhoff G D (1927) Dynamical Systems, American Mathematical Society, Providence, Rhode Island (reprint 1966).

Engel E M R A (1992) A Road to Randomness in Physical Systems, Springer, Berlin.

Herrmann D (1994) Algorithmen für Chaos und Fraktale, Addison-Wesley, Bonn.

Hilborn R C (1994) Chaos and Nonlinear Dynamics, Oxford Univ. Press.

Hopf E (1937) Ergodentheorie, Springer, Berlin (reprint 1970).

Jackson E A (1990) Perspectives of Nonlinear Dynamics, 2 vols., Cambridge Univ. Press.

Kac M (1959) Statistical Independence in Probability, Analysis and Number Theory, Wiley, New York.

Kolmogorov A N (1988) Mathematics — Science and Profession (in Russian), Nauka, Moscow.

Korsch H J and Jodl H J (1994) Chaos: A Program Collection for the PC, Springer, Berlin.

Moritz H (1995) Science, Mind and the Universe: An Introduction to Natural Philosophy, Wichmann, Heidelberg.

Poincaré H (1908) Science et Méthode, Flammarion, Paris; reprint: Science and Method, Dover, New York.

Poincaré H (1912) Calcul des Probabilités, Gauthier–Villars, Paris.

Ruhla C (1992) The Physics of Chance, Oxford Univ. Press.

SCIENCE AND METHOD

BY HENRI POINCARÉ. TRANSLATED BY FRANCIS MAITLAND.

DOVER PUBLICATIONS, INC.

1908

IV.

CHANCE.

I.

"How can we venture to speak of the laws of chance? Is not chance the antithesis of all law?" It is thus that Bertrand expresses himself at the beginning of his "Calculus of Probabilities." Probability is the opposite of certainty; it is thus what we are ignorant of, and consequently it would seem to be what we cannot calculate. There is here at least an apparent contradiction, and one on which much has already been written.

To begin with, what is chance? The ancients distinguished between the phenomena which seemed to obey harmonious laws, established once for all, and those that they attributed to chance, which were those that could not be predicted because they were not subject to any law. In each domain the precise laws did not decide everything, they only marked the limits within which chance was allowed to move. In this conception, the word chance had a precise, objective meaning; what was chance for one was also chance for the other and even for the gods.

But this conception is not ours. We have become complete determinists, and even those who wish to reserve the right of human free will at least allow determinism to reign undisputed in the inorganic world. Every phenomenon, however trifling it be, has a cause, and a mind infinitely powerful and infinitely well-informed concerning the laws of nature could have foreseen it from the beginning of the ages. If a being with such a mind existed, we could play no game of chance with him; we should always lose.

For him, in fact, the word chance would have no meaning, or rather there would be no such thing as chance. That there is for us is only on account of our frailty and our ignorance. And even without going beyond our frail humanity, what is chance for the ignorant is no longer chance for the learned. Chance is only the measure of our ignorance. Fortuitous phenomena are, by definition, those whose laws we are ignorant of.

But is this definition very satisfactory? When the first Chaldean shepherds followed with their eyes the movements of the stars, they did not yet know the laws of astronomy, but would they have dreamed of saying that the stars move by chance? If a modern physicist is studying a new phenomenon, and if he discovers its law on Tuesday, would he have said on Monday that the phenomenon was fortuitous? But more than this, do we not often invoke what Bertrand calls the laws of chance in order to predict a phenomenon? For instance, in the kinetic theory of gases, we find the well-known laws of Mariotte and of Gay-Lussac, thanks to the hypothesis that the velocities of the gaseous molecules vary irregularly, that is to say, by chance.

The observable laws would be much less simple, say all the physicists, if the velocities were regulated by some simple elementary law, if the molecules were, as they say, *organized*, if they were subject to some discipline. It is thanks to chance—that is to say, thanks to our ignorance, that we can arrive at conclusions. Then if the word chance is merely synonymous with ignorance, what does this mean? Must we translate as follows?—

"You ask me to predict the phenomena that will be produced. If I had the misfortune to know the laws of these phenomena, I could not succeed except by inextricable calculations, and I should have to give up the attempt to answer you; but since I am fortunate enough to be ignorant of them, I will give you an answer at once. And, what is more extraordinary still, my answer will be right."

Chance, then, must be something more than the name we give to our ignorance. Among the phenomena whose causes we are ignorant of, we must distinguish between fortuitous phenomena, about which the calculation of probabilities will give us provisional information, and those that are not fortuitous, about which we can say nothing, so long as we have not determined the laws that govern them. And as regards the fortuitous phenomena themselves, it is clear that the information that the calculation of probabilities supplies will not cease to be true when the phenomena are better known.

The manager of a life insurance company does not know when each of the assured will die, but he relies upon the calculation of probabilities and on the law of large numbers, and he does not make a mistake, since he is able to pay dividends to his shareholders. These dividends would not vanish if a very far-sighted and very indiscreet doctor came, when once the policies were signed, and gave the manager information on the chances of life of the assured. The doctor would dissipate the ignorance of the manager, but he would have no effect upon the dividends, which are evidently not a result of that ignorance.

II.

In order to find the best definition of chance, we must examine some of the facts which it is agreed to regard as fortuitous, to which the calculation of probabilities seems to apply. We will then try to find their common characteristics.

We will select unstable equilibrium as our first example. If a cone is balanced on its point, we know very well that it will fall, but we do not know to which side; it seems that chance alone will decide. If the cone were perfectly symmetrical, if its axis were perfectly vertical, if it were subject to no other force but gravity, it would not fall at all. But the slightest defect of symmetry will make it lean slightly to one side or other, and as soon as it leans, be it ever so little, it will fall altogether to that side. Even if the symmetry is perfect, a very slight trepidation, or a breath of air, may make it incline a few seconds of arc, and that will be enough to determine its fall and even the direction of its fall, which will be that of the original inclination.

A very small cause which escapes our notice determines a considerable effect that we cannot fail to see, and then we say that that effect is due to chance. If we knew exactly the laws of nature and the situation of the universe at the initial moment, we could predict exactly the situation of that same universe at a succeeding moment. But, even if it were the case that the natural laws had no longer any secret for us, we could still only know the initial situation *approximately*. If that enabled us to predict the succeeding situation *with the same approximation*, that is all we require, and we should say that the phenomenon had been predicted, that it is governed by laws. But it is not always so; it may happen that small differences in the initial conditions produce very great ones in the final phenomena. A small error in the former will produce an enormous error in the latter. Prediction becomes impossible, and we have the fortuitous phenomenon.

Our second example will be very much like our first, and we will borrow it from meteorology. Why have meteorologists such difficulty in predicting the weather with any certainty? Why is it that showers and even storms seem to come by chance, so that many people think it quite natural to pray for rain or fine weather, though they would consider it ridiculous to ask for an eclipse by prayer? We see that great disturbances are generally produced in regions where the atmosphere is in unstable equilibrium. The meteorologists see very well that the equilibrium is unstable, that a cyclone will be formed somewhere, but exactly where they are not in a position to say; a tenth of a degree more or less at any given point, and the cyclone will burst here and not there, and extend its ravages over districts it would otherwise have spared. If they had been aware of this tenth of a degree, they could have known it beforehand, but the observations were neither sufficiently comprehensive nor sufficiently precise, and that is the reason why it all seems due to the intervention of chance. Here, again, we find the same contrast between a very trifling cause that is inappreciable to the observer, and considerable effects, that are sometimes terrible disasters.

Let us pass to another example, the distribution of the minor planets on the Zodiac. Their initial longitudes may have had some definite order, but their mean motions were different and they have been revolving for so long that we may say that practically they are distributed *by chance* throughout the Zodiac. Very small initial differences in their distances from the sun, or, what amounts to the same thing, in their mean motions, have resulted in enormous differences in their actual longitudes. A difference of a thousandth part of a second in the mean daily motion will have the effect of a second in three years, a degree in ten thousand years, a whole circumference in three or four millions of years, and what is that beside the time that has elapsed since the minor planets became detached from Laplace's nebula? Here, again, we have a small cause and a great effect, or better, small differences in the cause and great differences in the effect.

The game of roulette does not take us so far as it might appear from the preceding example. Imagine a needle that can be turned about a pivot on a dial divided into a hundred alternate red and black sections. If the needle stops at a red section we win; if not, we lose. Clearly, all depends on the initial impulse we give to the needle. I assume that the needle will make ten or twenty revolutions, but it will stop earlier or later according to the strength of the spin I have given it. Only a variation of a thousandth or a two-thousandth in the impulse is sufficient to determine whether my needle will stop at a black section or at the following section, which is red. These are differences that the muscular sense cannot appreciate, which would escape even more delicate instruments. It is, accordingly, impossible for me to predict what the needle I have just spun will do, and that is why my heart beats and I hope for everything from chance. The difference in the cause is imperceptible, and the difference in the effect is for me of the highest importance, since it affects my whole stake.

III.

In this connexion I wish to make a reflection that is somewhat foreign to my subject. Some years ago a certain philosopher said that the future was determined by the past, but not the past by the future; or, in other words, that from the knowledge of the present we could deduce that of the future but not that of the past; because, he said, one cause can produce only one effect, while the same effect can be produced by several different causes. It is obvious that no scientist can accept this conclusion. The laws of nature link the antecedent to the consequent in such a way that the antecedent is determined by the consequent just as much as the consequent is by the antecedent. But what can have been the origin of the philosopher's error? We know that, in virtue of Carnot's principle, physical phenomena are irreversible and that the world is tending towards uniformity. When two bodies of different temperatures are in conjunction, the warmer gives up heat to the colder, and accordingly we can predict that the temperatures will become equal. But once the temperatures have become equal, if we are asked about the previous state, what can we answer? We can certainly say that one of the bodies was hot and the other cold, but we cannot guess which of the two was formerly the warmer.

And yet in reality the temperatures never arrive at perfect equality. The difference between the temperatures only tends towards zero asymptotically. Accordingly there comes a moment when our thermometers are powerless to disclose it. But if we had thermometers a thousand or a hundred thousand times more sensitive, we should recognize that there is still a small difference, and that one of the bodies has remained a little warmer than the other, and then we should be able to state that this is the one which was formerly very much hotter than the other.

So we have, then, the reverse of what we found in the preceding examples, great differences in the cause and small differences in the effect. Flammarion once imagined an observer moving away from the earth at a velocity greater than that of light. For him time would have its sign changed, history would be reversed, and Waterloo would come before Austerlitz. Well, for this observer effects and causes would be inverted, unstable equilibrium would no longer be the exception; on account of the universal irreversibility, everything would seem to him to come out of a kind of chaos in unstable equilibrium, and the whole of nature would appear to him to be given up to chance.

IV.

We come now to other arguments, in which we shall see somewhat different characteristics appearing, and first let us take the kinetic theory of gases. How are we to picture a receptacle full of gas? Innumerable molecules, animated with great velocities, course through the receptacle in all directions; every moment they collide with the sides or else with one another, and these collisions take place under the most varied conditions. What strikes us most in this case is not the smallness of the causes, but their complexity. And yet the former element is still found here, and plays an important part. If a molecule deviated from its trajectory to left or right in a very small degree as compared with the radius of action of the gaseous molecules, it would avoid a collision, or would suffer it under different conditions, and that would alter the direction of its velocity after the collision perhaps by 90 or 180 degrees.

That is not all. It is enough, as we have just seen, that the molecule should deviate before the collision in an infinitely small degree, to make it deviate after the collision in a finite degree. Then, if the molecule suffers two successive collisions, it is enough that it should deviate before the first collision in a degree of infinite smallness of the second order, to make it deviate after the first collision in a degree of infinite smallness of the first order, and after the second collision in a finite degree. And the molecule will not suffer two collisions only, but a great number each second.

So that if the first collision multiplied the deviation by a very large number, A, after n collisions it will be multiplied by A^n. It will, therefore, have become very great, not only because A is large—that is to say, because small causes produce great effects—but because the exponent n is large, that is to say, because the collisions are very numerous and the causes very complex.

Let us pass to a second example. Why is it that in a shower the drops of rain appear to us to be distributed by chance? It is again because of the complexity of the causes which determine their formation. Ions have been distributed through the atmosphere; for a long time they have been subjected to constantly changing air currents; they have been involved in whirlwinds of very small dimensions, so that their final distribution has no longer any relation to their original distribution. Suddenly the temperature falls, the vapour condenses, and each of these ions becomes the centre of a raindrop. In order to know how these drops will be distributed and how many will fall on each stone of the pavement, it is not enough to know the original position of the ions, but we must calculate the effect of a thousand minute and capricious air currents.

It is the same thing again if we take grains of dust in suspension in water. The vessel is permeated by currents whose law we know nothing of except that it is very complicated. After a certain length of time the grains will be distributed by chance, that is to say uniformly, throughout the vessel, and this is entirely due to the complication of the currents If they obeyed some simple law—if, for instance the vessel were revolving and the currents revolved in circles about its axis—the case would be altered, for each grain would retain its original height and its original distance from the axis.

We should arrive at the same result by picturing the mixing of two liquids or of two fine powders. To take a rougher example, it is also what happens when a pack of cards is shuffled. At each shuffle the cards undergo a permutation similar to that studied in the theory of substitutions. What will be the resulting permutation? The probability that it will be any particular permutation (for instance, that which brings the card occupying the position $\phi(n)$ before the permutation into the position n), this probability, I say, depends on the habits of the player. But if the player shuffles the cards long enough, there will be a great number of successive permutations, and the final order which results will no longer be governed by anything but chance; I mean that all the possible orders will be equally probable. This result is due to the great number of successive permutations, that is to say, to the complexity of the phenomenon.

A final word on the theory of errors. It is a case in which the causes have complexity and multiplicity. How numerous are the traps to which the observer is exposed, even with the best instrument. He must take pains to look out for and avoid the most flagrant, those which give birth to systematic errors. But when he has eliminated these, admitting that he succeeds in so doing, there still remain many which, though small, may become dangerous by the accumulation of their effects. It is from these that accidental errors arise, and we attribute them to chance, because their causes are too complicated and too numerous. Here again we have only small causes, but each of them would only produce a small effect; it is by their union and their number that their effects become formidable.

V.

There is yet a third point of view, which is less important than the two former, on which I will not lay so much stress. When we are attempting to predict a fact and making an examination of the antecedents, we endeavour to enquire into the anterior situation. But we cannot do this for every part of the universe, and we are content with knowing what is going on in the neighbourhood of the place where the fact will occur, or what appears to have some connexion with the fact. Our enquiry cannot be complete, and we must know how to select. But we may happen to overlook circumstances which, at first sight, seemed completely foreign to the anticipated fact, to which we should never have dreamed of attributing any influence, which nevertheless, contrary to all anticipation, come to play an important part.

A man passes in the street on the way to his business. Some one familiar with his business could say what reason he had for starting at such an hour and why he went by such a street. On the roof a slater is at work. The contractor who employs him could, to a certain extent, predict what he will do. But the man has no thought for the slater, nor the slater for him; they seem to belong to two worlds completely foreign to one another. Nevertheless the slater drops a tile which kills the man, and we should have no hesitation in saying that this was chance.

Our frailty does not permit us to take in the whole universe, but forces us to cut it up in slices. We attempt to make this as little artificial as possible, and yet it happens, from time to time, that two of these slices react upon each other, and then the effects of this mutual action appear to us to be due to chance.

Is this a third way of conceiving of chance? Not always; in fact, in the majority of cases, we come back to the first or second. Each time that two worlds, generally foreign to one another, thus come to act upon each other, the laws of this reaction cannot fail to be very complex, and moreover a very small change in the initial conditions of the two worlds would have been enough to prevent the reaction from taking place. How very little it would have taken to make the man pass a moment later, or the slater drop his tile a moment earlier!

VI.

Nothing that has been said so far explains why chance is obedient to laws. Is the fact that the causes are small, or that they are complex, sufficient to enable us to predict, if not what the effects will be *in each case*, at least what they will be *on the average?* In order to answer this question, it will be best to return to some of the examples quoted above.

I will begin with that of roulette. I said that the point where the needle stops will depend on the initial impulse given it. What is the probability that this impulse will be of any particular strength? I do not know, but it is difficult not to admit that this probability is represented by a continuous analytical function. The probability that the impulse will be comprised between α and $\alpha+\epsilon$ will, then, clearly be equal to the probability that it will be comprised between $\alpha+\epsilon$ and $\alpha+2\epsilon$, *provided that* ϵ *is very small.* This is a property common to all analytical functions. Small variations of the function are proportional to small variations of the variable.

But we have assumed that a very small variation in the impulse is sufficient to change the colour of the section opposite which the needle finally stops. From α to $\alpha+\epsilon$ is red, from $\alpha+\epsilon$ to $\alpha+2\epsilon$ is black. The probability of each red section is accordingly the same as that of the succeeding black section, and consequently the total probability of red is equal to the total probability of black.

The datum in the case is the analytical function which represents the probability of a particular initial impulse. But the theorem remains true, whatever this datum may be, because it depends on a property common to all analytical functions. From this it results finally that we have no longer any need of the datum.

What has just been said of the case of roulette applies also to the example of the minor planets. The Zodiac may be regarded as an immense roulette board on which the Creator has thrown a very great number of small balls, to which he has imparted different initial impulses, varying, however, according to some sort of law. Their actual distribution is uniform and independent of that law, for the same reason as in the preceding case. Thus we see why phenomena obey the laws of chance when small differences in the causes are sufficient to produce great differences in the effects. The probabilities of these small differences can then be regarded as proportional to the differences themselves, just because these differences are small, and small increases of a continuous function are proportional to those of the variable.

Let us pass to a totally different example, in which the complexity of the causes is the principal factor. I imagine a card-player shuffling a pack of cards. At each shuffle he changes the order of the cards, and he may change it in various ways. Let us take three cards only in order to simplify the explanation. The cards which, before the shuffle, occupied the positions 1 2 3 respectively may, after the shuffle, occupy the positions

$$1\,2\,3,\ 2\,3\,1,\ 3\,1\,2,\ 3\,2\,1,\ 1\,3\,2,\ 2\,1\,3.$$

Each of these six hypotheses is possible, and their probabilities are respectively

$$p_1,\ p_2,\ p_3,\ p_4,\ p_5,\ p_6.$$

The sum of these six numbers is equal to 1, but that is all we know about them. The six probabilities naturally depend upon the player's habits, which we do not know.

At the second shuffle the process is repeated, and under the same conditions. I mean, for instance, that p_4 always represents the probability that the three cards which occupied the positions 1 2 3 after the n^{th} shuffle and before the $n+1^{th}$, will occupy the positions 3 2 1 after the $n+1^{th}$ shuffle. And this remains true, whatever the number n may be, since the player's habits and his method of shuffling remain the same.

But if the number of shuffles is very large, the cards which occupied the positions 123 before the first shuffle may, after the last shuffle, occupy the positions

123, 231, 312, 321, 132, 213,

and the probability of each of these six hypotheses is clearly the same and equal to $\frac{1}{6}$; and this is true whatever be the numbers $p_1 \ldots p_6$, which we do not know. The great number of shuffles, that is to say, the complexity of the causes, has produced uniformity.

This would apply without change if there were more than three cards, but even with three the demonstration would be complicated, so I will content myself with giving it for two cards only. We have now only two hypotheses

12, 21,

with the probabilities p_1 and $p_2 = 1 - p_1$. Assume that there are n shuffles, and that I win a shilling if the cards are finally in the initial order, and that I lose one if they are finally reversed. Then my mathematical expectation will be

$$(p_1 - p_2)^n$$

The difference $p_1 - p_2$ is certainly smaller than 1, so that if n is very large, the value of my expectation will be nothing, and we do not require to know p_1 and p_2 to know that the game is fair.

Nevertheless there would be an exception if one of the numbers p_1 and p_2 was equal to 1 and the other to nothing. *It would then hold good no longer, because our original hypotheses would be too simple.*

What we have just seen applies not only to the mixing of cards, but to all mixing, to that of powders and liquids, and even to that of the gaseous molecules in the kinetic theory of gases. To return to this theory, let us imagine for a moment a gas whose molecules cannot collide mutually, but can be deviated by collisions with the sides of the vessel in which the gas is enclosed. If the form of the vessel is sufficiently complicated, it will not be long before the distribution of the molecules and that of their velocities become uniform. This will not happen if the vessel is spherical, or if it has the form of a rectangular parallelepiped. And why not? Because in the former case the distance of any particular trajectory from the centre remains constant, and in the latter case we have the absolute value of the angle of each trajectory with the sides of the parallelepiped.

Thus we see what we must understand by conditions that are *too simple*. They are conditions which preserve something of the original state as an invariable. Are the differential equations of the problem too simple to enable us to apply the laws of chance? This question appears at first sight devoid of any precise meaning, but we know now what it means. They are too simple if something is preserved, if they admit a uniform integral. If something of the initial conditions remains unchanged, it is clear that the final situation can no longer be independent of the initial situation.

We come, lastly, to the theory of errors. We are ignorant of what accidental errors are due to, and it is just because of this ignorance that we know they will obey Gauss's law. Such is the paradox. It is explained in somewhat the same way as the preceding cases. We only need to know one thing—that the errors are very numerous, that they are very small, and that each of them can be equally well negative or positive. What is the curve of probability of each of them? We do not know, but only assume that it is symmetrical. We can then show that the resultant error will follow Gauss's law, and this resultant law is independent of the particular laws which we do not know. Here again the simplicity of the result actually owes its existence to the complication of the data.

VII.

But we have not come to the end of paradoxes. I recalled just above Flammarion's fiction of the man who travels faster than light, for whom time has its sign changed. I said that for him all phenomena would seem to be due to chance. This is true from a certain point of view, and yet, at any given moment, all these phenomena would not be distributed in conformity with the laws of chance, since they would be just as they are for us, who, seeing them unfolded harmoniously and not emerging from a primitive chaos, do not look upon them as governed by chance.

What does this mean? For Flammarion's imaginary Lumen, small causes seem to produce great effects; why, then, do things not happen as they do for us when we think we see great effects due to small causes? Is not the same reasoning applicable to his case?

Let us return to this reasoning. When small differences in the causes produce great differences in the effects, why are the effects distributed according to the laws of chance? Suppose a difference of an inch in the cause produces a difference of a mile in the effect. If I am to win in case the effect corresponds with a mile bearing an even number, my probability of winning will be ½. Why is this? Because, in order that it should be so, the cause must correspond with an inch bearing an even number. Now, according to all appearance, the probability that the cause will vary between certain limits is proportional to the distance of those limits, provided that distance is very small. If this hypothesis be not admitted, there would no longer be any means of representing the probability by a continuous function.

Now what will happen when great causes produce small effects? This is the case in which we shall not attribute the phenomenon to chance, and in which Lumen, on the contrary, would attribute it to chance. A difference of a mile in the cause corresponds to a difference of an inch in the effect. Will the probability that the cause will be comprised between two limits n miles apart still be proportional to n? We have no reason to suppose it, since this distance of n miles is great. But the probability that the effect will be comprised between two limits n inches apart will be precisely the same, and accordingly it will not be proportional to n, and that notwithstanding the fact that this distance of n inches is small. There is, then, no means of representing the law of probability of the effects by a continuous curve. I do not mean to say that the curve may not remain continuous in the *analytical* sense of the word. To *infinitely small* variations of the abscissa there will correspond infinitely small variations of the ordinate. But *practically* it would not be continuous, since to *very small* variations of the abscissa there would not correspond very small variations of the ordinate. It would become impossible to trace the curve with an ordinary pencil: that is what I mean.

What conclusion are we then to draw? Lumen has no right to say that the probability of the cause (that of *his* cause, which is our effect) must necessarily be represented by a continuous function. But if that be so, why have we the right? It is because that state of unstable equilibrium that I spoke of just now as initial, is itself only the termination of a long anterior history. In the course of this history complex causes have been at work, and they have been at work for a long time. They have contributed to bring about the mixture of the elements, and they have tended to make everything uniform, at least in a small space. They have rounded off the corners, levelled the mountains, and filled up the valleys. However capricious and irregular the original curve they have been given, they have worked so much to regularize it that they will finally give us a continuous curve, and that is why we can quite confidently admit its continuity.

Lumen would not have the same reasons for drawing this conclusion. For him complex causes would not appear as agents of regularity and of levelling; on the contrary, they would only create differentiation and inequality. He would see a more and more varied world emerge from a sort of primitive chaos. The changes he would observe would be for him unforeseen and impossible to foresee. They would seem to him due to some caprice, but that caprice would not be at all the same as our chance, since it would not be amenable to any law, while our chance has its own laws. All these points would require a much longer development, which would help us perhaps to a better comprehension of the irreversibility of the universe.

VIII.

We have attempted to define chance, and it would be well now to ask ourselves a question. Has chance, thus defined so far as it can be, an objective character?

We may well ask it. I have spoken of very small or very complex causes, but may not what is very small for one be great for another, and may not what seems very complex to one appear simple to another? I have already given a partial answer, since I stated above most precisely the case in which differential equations become too simple for the laws of chance to remain applicable. But it would be well to examine the thing somewhat more closely, for there are still other points of view we may take.

What is the meaning of the word small? To understand it, we have only to refer to what has been said above. A difference is very small, an interval is small, when within the limits of that interval the probability remains appreciably constant. Why can that probability be regarded as constant in a small interval? It is because we admit that the law of probability is represented by a continuous curve, not only continuous in the analytical sense of the word, but *practically* continuous, as I explained above. This means not only that it will present no absolute hiatus, but also that it will have no projections or depressions too acute or too much accentuated.

What gives us the right to make this hypothesis?

As I said above, it is because, from the beginning of the ages, there are complex causes that never cease to operate in the same direction, which cause the world to tend constantly towards uniformity without the possibility of ever going back. It is these causes which, little by little, have levelled the projections and filled up the depressions, and it is for this reason that our curves of probability present none but gentle undulations. In millions and millions of centuries we shall have progressed another step towards uniformity, and these undulations will be ten times more gentle still. The radius of mean curvature of our curve will have become ten times longer. And then a length that to-day does not seem to us very small, because an arc of such a length cannot be regarded as rectilineal, will at that period be properly qualified as very small, since the curvature will have become ten times less, and an arc of such a length will not differ appreciably from a straight line.

Thus the word very small remains relative, but it is not relative to this man or that, it is relative to the actual state of the world. It will change its meaning when the world becomes more uniform and all things are still more mixed. But then, no doubt, men will no longer be able to live, but will have to make way for other beings, shall I say much smaller or much larger? So that our criterion, remaining true for all men, retains an objective meaning.

And, further, what is the meaning of the word very complex? I have already given one solution, that which I referred to again at the beginning of this section; but there are others. Complex causes, I have said, produce a more and more intimate mixture, but how long will it be before this mixture satisfies us? When shall we have accumulated enough complications? When will the cards be sufficiently shuffled? If we mix two powders, one blue and the other white, there comes a time when the colour of the mixture appears uniform. This is on account of the infirmity of our senses; it would be uniform for the long-sighted, obliged to look at it from a distance, when it would not yet be so for the short-sighted. Even when it had become uniform for all sights, we could still set back the limit by employing instruments. There is no possibility that any man will ever distinguish the infinite variety that is hidden under the uniform appearance of a gas, if the kinetic theory is true. Nevertheless, if we adopt Gouy's ideas on the Brownian movement, does not the microscope seem to be on the point of showing us something analogous?

This new criterion is thus relative like the first, and if it preserves an objective character, it is because all men have about the same senses, the power of their instruments is limited, and, moreover, they only make use of them occasionally.

IX.

It is the same in the moral sciences, and particularly in history. The historian is obliged to make a selection of the events in the period he is studying, and he only recounts those that seem to him the most important. Thus he contents himself with relating the most considerable events of the 16th century, for instance, and similarly the most remarkable facts of the 17th century. If the former are sufficient to explain the latter, we say that these latter conform to the laws of history. But if a great event of the 17th century owes its cause to a small fact of the 16th century that no history reports and that every one has neglected, then we say that this event is due to chance, and so the word has the same sense as in the physical sciences; it means that small causes have produced great effects.

The greatest chance is the birth of a great man. It is only by chance that the meeting occurs of two genital cells of different sex that contain precisely, each on its side, the mysterious elements whose mutual reaction is destined to produce genius. It will be readily admitted that these elements must be rare, and that their meeting is still rarer. How little it would have taken to make the spermatozoid which carried them deviate from its course. It would have been enough to deflect it a hundredth part of a inch, and Napoleon would not have been born and the destinies of a continent would have been changed. No example can give a better comprehension of the true character of chance.

One word more about the paradoxes to which the application of the calculation of probabilities to the moral sciences has given rise. It has been demonstrated that no parliament would ever contain a single member of the opposition, or at least that such an event would be so improbable that it would be quite safe to bet against it, and to bet a million to one. Condorcet attempted to calculate how many jurymen it would require to make a miscarriage of justice practically impossible. If we used the results of this calculation, we should certainly be exposed to the same disillusionment as by betting on the strength of the calculation that the opposition would never have a single representative.

The laws of chance do not apply to these questions. If justice does not always decide on good grounds, it does not make so much use as is generally supposed of Bridoye's method. This is perhaps unfortunate, since, if it did, Condorcet's method would protect us against miscarriages.

What does this mean? We are tempted to attribute facts of this nature to chance because their causes are obscure, but this is not true chance. The causes are unknown to us, it is true, and they are even complex; but they are not sufficiently complex, since they preserve something, and we have seen that this is the distinguishing mark of "too simple" causes. When men are brought together, they no longer decide by chance and independently of each other, but react upon one another. Many causes come into action, they trouble the men and draw them this way and that, but there is one thing they cannot destroy, the habits they have of Panurge's sheep. And it is this that is preserved.

X.

The application of the calculation of probabilities to the exact sciences also involves many difficulties. Why are the decimals of a table of logarithms or of the number π distributed in accordance with the laws of chance? I have elsewhere studied the question in regard to logarithms, and there it is easy. It is clear that a small difference in the argument will give a small difference in the logarithm, but a great difference in the sixth decimal of the logarithm. We still find the same criterion.

But as regards the number π the question presents more difficulties, and for the moment I have no satisfactory explanation to give.

There are many other questions that might be raised, if I wished to attack them before answering the one I have more especially set myself. When we arrive at a simple result, when, for instance, we find a round number, we say that such a result cannot be due to chance, and we seek for a non-fortuitous cause to explain it. And in fact there is only a very slight likelihood that, out of 10,000 numbers, chance will give us a round number, the number 10,000 for instance; there is only one chance in 10,000. But neither is there more than one chance in 10,000 that it will give us any other particular number, and yet this result does not astonish us, and we feel no hesitation about attributing it to chance, and that merely because it is less striking.

Is this a simple illusion on our part, or are there cases in which this view is legitimate? We must hope so, for otherwise all science would be impossible. When we wish to check a hypothesis, what do we do? We cannot verify all its consequences, since they are infinite in number. We content ourselves with verifying a few, and, if we succeed, we declare that the hypothesis is confirmed, for so much success could not be due to chance. It is always at bottom the same reasoning.

I cannot justify it here completely, it would take me too long, but I can say at least this. We find ourselves faced by two hypotheses, either a simple cause or else that assemblage of complex causes we call chance. We find it natural to admit that the former must produce a simple result, and then, if we arrive at this simple result, the round number for instance, it appears to us more reasonable to attribute it to the simple cause, which was almost certain to give it us, than to chance, which could only give it us once in 10,000 times. It will not be the same if we arrive at a result that is not simple. It is true that chance also will not give it more than once in 10,000 times, but the simple cause has no greater chance of producing it.

Helmut Moritz to Lecture Notes of Prof. H. MORITZ

Science, Mind and the Universe

An Introduction to Natural Philosophy

WICHMANN 1995

Chapter 3

Physics

3.1 Classical mechanics and determinism

Nature and Nature's Laws lay hid in Night:
God said, Let Newton be! and All was Light.
Alexander Pope

The great Greek philosopher Aristotle (384–322 B.C.) believed that the *velocity* of a body is proportional to the force to which it was subjected. Ordinary experience seems to confirm this view. A horse carriage moves the faster, the stronger the horses are. A body lying on the floor does not move unless some force is exerted to drag it along.

Only Galileo Galilei (1564–1642) recognized that matters are not so simple. A body lying on a very smooth and plane ice surface will continue to move with constant velocity and in a constant direction even if the initial force has ceased to act. To be sure, this body will gradually slow down and finally stop, but the cause is *friction*. If there is no friction, the movement will be continuous and will never come to a stop. A space ship in intergalactic space will forever move with constant speed along a straight line after the rocket engines have been shut off. Thus Aristotle and common sense have been deceived by friction.

The correct law of motion in the absence of friction was discovered by Isaac Newton (1642–1727). It has the form

$$m\ddot{\mathbf{x}} = \mathbf{F} \quad . \tag{3.1}$$

Here m denotes the mass, and $\mathbf{F}$ is the force. The position vector is

$$\mathbf{x} = [x, y, z] \quad , \tag{3.2}$$

the velocity vector is its time derivative:

$$\dot{\mathbf{x}} \equiv \frac{d\mathbf{x}}{dt} = [\dot{x}, \dot{y}, \dot{z}] \quad , \tag{3.3}$$

and the acceleration is the second derivative:

$$\ddot{\mathbf{x}} \equiv \frac{d^2\mathbf{x}}{dt^2} = [\ddot{x}, \ddot{y}, \ddot{z}] \quad . \tag{3.4}$$

Thus *Newton's law of motion* (3.1) says that the *acceleration* is proportional to the force, and not the velocity as Aristotle thought.

In order to fully define the movement, in addition to the *differential equation* (3.1) we need *initial conditions*: at a certain instant $t = t_0$, the position and velocity:

$$\mathbf{x}_0 = \mathbf{x}(t_0) \quad , \qquad \dot{\mathbf{x}}_0 = \dot{\mathbf{x}}(t_0) \tag{3.5}$$

must be given.

Assume motion under no force, $\mathbf{F} = 0$. Then (3.1) gives

$$\ddot{\mathbf{x}} = 0 \quad . \tag{3.6}$$

The solution of this differential equation is

$$\mathbf{x} = \mathbf{a}t + \mathbf{b} \quad , \tag{3.7}$$

where the constant vectors $\mathbf{a}$ and $\mathbf{b}$ serve as integration constants. To understand this, differentiate (3.7) twice:

$$\dot{\mathbf{x}} = \mathbf{a} \quad , \tag{3.8}$$
$$\ddot{\mathbf{x}} = 0 \quad . \tag{3.9}$$

Thus (3.6) is satisfied, what was to be shown. If we put $t = 0$ in (3.7) and (3.8), we get

$$\mathbf{a} = \dot{\mathbf{x}}_0 \quad , \qquad \mathbf{b} = \mathbf{x}_0 \quad . \tag{3.10}$$

Taking for the initial instant $t_0 = 0$, we thus have a very instructive interpretation of the integration constants $\mathbf{a}$ and $\mathbf{b}$: they are nothing else than the initial conditions (3.5).

Newton's law of gravitation. Besides Newton's law of motion (3.1), we also have his law of gravitation:

$$F = G\frac{m_1 m_2}{l^2} \quad . \tag{3.11}$$

Two point masses (Fig. 3.1) attract each other with a force $\mathbf{F}$ of magnitude F, proportional to the masses m_1 and m_2, and inversely proportional to the square of their distance l; this is the famous ***inverse square law.*** Here G denotes a universal constant, the ***gravitational constant.*** We also have the equality of ***action and reaction***: the two forces $\mathbf{F}_1$ and $\mathbf{F}_2$ in Fig. 3.1 are equal in magnitude and opposite in direction. The magnitude of both $\mathbf{F}_1$ and $\mathbf{F}_2$ is given by (3.11).

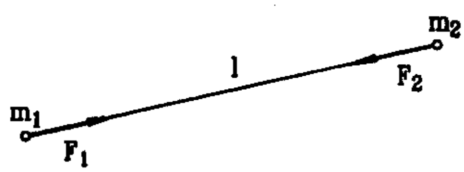

Figure 3.1: Illustrating the law of gravitation

If Newton's law of gravitation (3.11) is used in the equation of motion (3.1), then this differential equation, on integration, gives the Kepler ellipses, along which the planets move around the Sun.

Principles of mechanics. If the motion is subject to constraints, the simple Newtonian equation of motion is no longer applicable. For instance, frictionless motion of a particle constrained to move along a curved surface cannot be along a straight line, even if there is no external force, $\mathbf{F} = 0$. The "straightest" curve on a surface is a *geodesic*, representing the shortest line between two points that wholly lies in the surface. If the surface is a sphere, then the geodesic is a great circle. Now it can be shown that *frictionless and forceless motion along a surface really is motion with constant velocity along a geodesic.* Even this simple but important case is not covered by (3.1).

So for motion on a surface Newton's equation (3.1) is not satisfied, that is,

$$m\ddot{\mathbf{x}} - \mathbf{F} \neq 0 \quad . \tag{3.12}$$

If the left-hand side cannot be zero, then let us try at least to make it as small as possible:

$$(m\ddot{\mathbf{x}} - \mathbf{F})^2 \Longrightarrow \text{minimum} \quad , \tag{3.13}$$

subject to the given conditions, for instance, motion on a surface. This is *Gauss' principle of least constraint.*

It is in full analogy to the *principle of least squares* discussed in sec. 2.6, eq. (2.35) on p. 65. In fact, (2.36) says that

$$\mathbf{A}\,\mathbf{x} - \mathbf{l} \neq 0 \quad , \tag{3.14}$$

and (2.39) is equivalent to

$$(\mathbf{A}\,\mathbf{x} - \mathbf{l})^2 \Longrightarrow \text{minimum} \quad . \tag{3.15}$$

The analogy between (3.12) and (3.13), on the one hand, and (3.14) and (3.15), on the other hand, is obvious.

Thus it is not surprising that both principles are due to Gauss, who also recognized the deep analogy between them.

It may be shown that Gauss' principle applied to a free particle on a surface, does give geodesic motion. For the sphere, motion along a great circle is obvious, cf. Fig. 3.2.

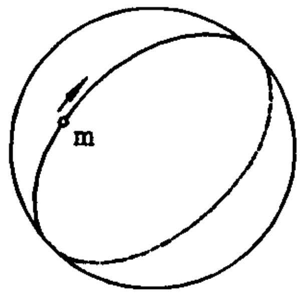

Figure 3.2: A free particle describes a geodesic on a sphere

For many other simple and complicated cases, Newton's elementary law (3.1) does not directly apply. A pertinent example is the rotation of a rigid body, because Newton's equations are essentially valid for point masses only and do not apply to rotation. With respect to orbital motion about the Sun, the planets may be considered point masses, but Earth rotation must be treated in a different way.

A number of other principles, more general than Newton's laws, were proposed in the 18th century by d'Alembert, Lagrange and others. This is subject of *analytical dynamics*, of which a non–specialist account can be found in (Lindsay and Margenau 1957, Chapter III). We have briefly considered only Gauss' principle and shall now outline Hamilton's method.

Hamilton's equations. The Newton equation (3.1) is in reality a system of *three* ordinary differential equations *of second order*:

$$\begin{aligned} m\ddot{x}_1 &= F_1(x_1, x_2, x_3) \quad , \\ m\ddot{x}_2 &= F_2(x_1, x_2, x_3) \quad , \\ m\ddot{x}_3 &= F_3(x_1, x_2, x_3) \quad , \end{aligned} \tag{3.16}$$

$x_1 = x$, $x_2 = y$, $x_3 = z$ denoting the Cartesian coordinates which are the components of the position vector $\mathbf{x}$, and similarly F_1, F_2, F_3 for the force vector $\mathbf{F}$.

Now we introduce the auxiliary quantities

$$p_1 = m\dot{x}_1 , \quad p_2 = m\dot{x}_2 , \quad p_3 = m\dot{x}_3 \tag{3.17}$$

called *momenta*. The coordinates x_1, x_2, x_3 are now denoted by q_1, q_2, q_3. Then (3.17) and (3.16) become with $i = 1, 2, 3$:

$$\begin{aligned} \dot{q}_i &= \frac{1}{m} p_i \quad , \\ \dot{p}_i &= F_i(q_1, q_2, q_3) \quad . \end{aligned} \tag{3.18}$$

Thus we have reduced the *three* differential equations (3.16) of *second* order by $3 + 3 =$ *six* differential equations of *first* order.

This method is standard in the theory of differential equations and not particularly enlightening.

What is significant, however, is the fact that William Hamilton (1788–1856) was able to bring (3.18) to the form

$$\begin{aligned} \dot{q}_i &= \frac{\partial H}{\partial p_i} \quad , \\ \dot{p}_i &= -\frac{\partial H}{\partial q_i} \quad , \end{aligned} \tag{3.19}$$

with *one* function H only, instead of the 3 functions F_i! Because of their importance, they are called the *canonical equations* of mechanics, and H is known as Hamilton's function or, briefly, as *Hamiltonian*. By the way, H is simply the sum of kinetic and potential energy. Any quantities p_i and q_i satisfying (3.19) are called *canonically conjugate variables.*

The true importance of the Hamiltonian equations (3.19), however, is the fact that q_i need not be Cartesian coordinates but can be any *generalized coordinates* (parameters), and i need not be restricted to 1, 2, 3 but can assume so many values as we need parameters to fully describe the dynamical system. For instance, for a rotating rigid body we need 6 parameters q_i ($i = 1, 2, 3, 4, 5, 6$): three translations (along the x, y, z axes) and three rotations (e.g., around the same axes). If we have r particles, then we need $3r$ parameters q_i: 3 for each particle.

Let us assume that we have n generalized coordinates q_i. Then we have $2n$ differential equations (3.19), and we can solve them uniquely provided we have the $2n$ *initial values* q_i and p_i at time $t = t_0$.

Laplace's demon.

An intelligent being which, for some given moment of time, knew all the forces by which nature is driven, and the relative position of the objects by which it is composed (provided the being's intelligence were so vast as to be able to analyze all the data), would be able to comprise, in a single formula, the movements of the largest bodies in the universe and those of the lightest atom: nothing would be uncertain to it, and both the future and the past would be present to its eyes. The human mind offers in the perfection which it has been able to give to astronomy, a feeble inkling of such an intelligence.

This impressive statement was given by Pierre Simon de Laplace (1749–1827); the "intelligent being" has become famous as "Laplace's demon".

This is the classical expression of *causality* or *determinism*: given the equations of motion and the initial conditions at $t = t_0$, the state of the system is exactly known at all earlier ($t < t_0$) and all later ($t > t_0$) times. Determinism reigned supreme until about 1925, when quantum theory started thoroughly to shake it (sec. 3.5).

Recently, however, determinism has come under attack even from its very stronghold, classical mechanics. This has been achieved by the theory of chaotic systems (sec. 3.2).

The principle of least action. Instead of differential equations, classical mechanics can also be expressed by an *integral* minimum principle of form

$$\int_A^B L\,dt \Longrightarrow \text{minimum} \tag{3.20}$$

where an integral (the "action") of a function L is to be minimized. The *Lagrangian* L is related to the energy and also to the Hamiltonian H in a way which is not necessary for the present argument. Least-action principles have been given by several scientists starting with Pierre Louis de Maupertuis (1698–1759) and Leonhard Euler (1707–1783).

From the *integral* principle (3.20) it is possible uniquely to derive the *differential* equations (3.19). This is of considerable philosophical importance, for the following reasons.

An integral principle (3.20), minimizing (or maximizing) some "overall" quantity, has been interpreted as expressing a tendency of nature towards perfection, attaining some ideal: maximum or optimum sounds better than minimum, but is essentially the same thing. It thus expresses a *finalist tendency*, a "causa finalis" in the sense of Aristotle, cf. sec. 5.4. Such finalism occurs especially in biology (sec. 4.1). It has been opposed to the causal determinism as exemplified by the differential equations of classical mechanics.

The deduction of the *deterministic* equations (3.19) from the *finalistic* integral (3.20) shows that both principles can coexist peacefully: the principle (3.20), so to speak, creates its own differential equations (3.19).

In a similar way we shall see in sec. 4.1 that a thermostat, governed by a "finalistic" principle of producing a desired temperature, will "generate" its own physical "deterministic" differential equations that help achieve the goal.

Thus causality, characteristic for classical mechanics, and finalism, considered typical for biology, are far less incompatible as they first appear, cf. also (Thom 1975, sec. 12.1.A).

"Causality", so to speak, is the answer to the question "For which reason?", whereas "finality" answers the question "For which purpose?".

The basic results of the present section will also be needed to discuss geodesic motion in general relativity (sec. 3.4) and a generalization of Hamiltonian methods to quantum theory (sec. 3.5).

But also taken in itself, classical mechanics has an incredibly rich structure. It comprises:

- causality: basic property;
- chaos: sec. 3.2;
- final causation: just discussed;
- constraints: eq. (3.13); and even
- "software laws" in a rudimentary form: as initial conditions (see also sec. 4.5).

Ideas will be needed rather than formulas, so the reader need not understand all mathematical details. Interested readers may consult any textbook on theoretical physics; particularly suited for the present purpose is the treatment in (Lindsay and Margenau 1957, Chapter III). We also mention (Margenau 1950) which is less mathematical and more philosophical and which is still a classic.

3.2 Deterministic chaos

In the beginning ... there was Poincaré.
E. Atlee Jackson

The deterministic paradise of classical mechanics, over which Laplace's demon (sec. 3.1) exerted a rigid but essentially benevolent, orderly, and stable regime, began to show, on closer inspection, some strange and irritating features.

The application of mechanics to gases and fluids consisting of an enormous number of particles (molecules) led to the statistical theory of heat. Heat was explained as the random and irregular, more or less violent motion of these particles. In view of the enormous number of these particles, it is practically impossible to describe the trajectory of every particle by Newton's laws (even assumed that this would be theoretically possible). Instead, these particles were treated statistically, which led to *statistical mechanics* or *statistical thermodynamics*, created by Josia Willard Gibbs (1839–1903), Ludwig Boltzmann (1844–1906) and others. A brilliant success was the derivation of the basic equations of thermodynamics from the principles of classical mechanics combined with statistical considerations. *Temperature* was explained in terms of the average kinetic energy of the molecules; it is the higher, the greater the average velocity of the particles is. The important concept of *entropy* was introduced, and Boltzmann found his famous equation, formula (4.3) of sec. 4.3.

But here a problem arises. The equations of classical mechanics are *time-reversible*. This means that these equations retain their form on replacing time t by $-t$. On the other hand, the equations of thermodynamics are typically *irreversible*: the entropy in a physical system always increases, see eq. (4.4). This contradiction must be due to the introduction of statistics, either because of the enormous amount of particles, or because of the incredibly complicated, "chaotic", shape of the trajectories of the particles (or both). These controverses, in which already Boltzmann was involved, led to very important advances in physics, mathematics, and probability theory (sec. 3.3), known by the name of *ergodic theory*.

The French mathematician Henri Poincaré (1854–1912) found already in 1890 that even relatively "simple" *nonlinear* dynamical problems in astronomy etc. may admit extremely complicated, irregular, even "chaotic" trajectories. In his classical work "Les Méthodes nouvelles de la Mécanique céleste" (1899) vol. III, p. 389 he wrote:

Imagine the figure formed by these two curves and their infinitely many intersections ...; these intersections form a kind of meshwork, tissue, or infinitely dense network ... One is struck by the complexity of this figure which I do not even attempt to draw. Nothing is better suited to give us an idea of the complexity of the three-body problem and in general of all the problems of dynamics in which there is no uniform integral [of the motion] ...

The modern theory of *general nonlinear dynamical systems* is considered to start with Poincaré's work. The subject then lay relatively dormant, known only to a few specialists, until 1954 when the famous Russian mathematician Andrei Kolmogorov (1903–1987) and his younger colleague Vladimir Arnold started with a general and systematic treatment of such strange trajectories. In 1963 there followed an independent paper on an application to meteorology by the American Edward Lorenz. Then the subject exploded. Currently it is probably the most popular subject of mathematics, known to a broad general public.

Let me try to explain what Lorenz did. He took the equations of mathematical *weather prediction*, simplified them and studied the solution numerically with the help of a computer. These solutions proved to be extremely *unstable*: two solutions with almost identical initial conditions started to diverge wildly (Fig. 3.3). Since the data of meteorology are unavoidably insufficient and inaccurate, the initial conditions are not exactly known; small deviations result in completely different behavior. This is the reason why it is hardly meaningful to make detailed weather predictions more than a few days ahead. (In astronomy, predictions are good for tens or even hundreds of years, in spite of Poincaré ...)

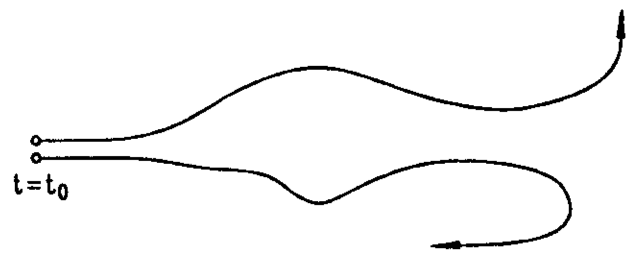

Figure 3.3: Two unstable trajectories

Let us repeat:

stability: small causes produce small effects;
instability: small causes produce large effects.

Classical causality implicitly presupposes stability. Stability is the environment in which Laplace's demon thrives.

Unstable systems are mathematically described always by *nonlinear* differential equations. Therefore, as we have already mentioned, mathematicians speak of general nonlinear dynamical systems. (Popularly speaking, the difference between "linear" and "nonlinear" is essentially the difference between a straight and a curved line; the function $y = 2 + 3x$ is linear, whereas the functions $y = x^2$ and $y = \sin x$ are nonlinear.) Unstable nonlinear dynamical systems are nowadays widely known by the name of *chaos theory.*

We distinguish between *conservative* dynamic systems for which the total energy is conserved (e.g., those described by Hamiltonian equations (3.19)), and *dissipative* systems for which part of the energy is dissipated as heat, e.g., through friction.

The nonlinear systems of *celestial mechanics* as investi gated by Poincaré, Kolmogorov, and Arnold are *conservative.* The *meteorologic systems* studied by Lorenz are *dissipative*, because the atmosphere constantly receives energy from the sun and radiates it again into outer space: otherwise "global warming" would be very rapid indeed. The name, *chaotic systems*, is particularly appropriate for meteorological and similar dynamic systems.

Chaos theory is an outstanding example of a theory as an instrument for discovery, a "searchlight": now chaotic phenomena are found everywhere, from clouds to earthquakes, and from turbulent mountain streams to human heartbeats. Deterministic chaos, so to speak, is an example of *chaos out of order.* There is also an emergence of *order out of chaos*; cf. the derivation of thermodynamics from statistical mechanics and sec. 3.3. In fact, both cases are closely interrelated and related also to the production of *order out of order*, cf. sec. 4.3 (p. 182).

The historian of science, W. Schröder, tells me that the well-known German meteorologist H. Ertel has found instability as the reason for the impossibility of weather prediction beyond a few days already in 1941. Ertel must therefore be considered a predecessor of Edward Lorenz in meteorological chaos. This is also true already for Poincarè (1908), as the quotation in sec. 6.3 (p. 243) shows.

Suggested additional reading. There is an incredible amount of books and papers on chaos theory. An advantage of its popularity is the fact that there are outstanding presentations for the general public, of high level but without formulas. An extremely readable introduction is (Gleick 1988); Stewart (1990) is a fascinating presentation of all the details but without formulas; an authoritative and very readable introduction is (Lorenz 1993); and Abraham and Shaw (1984) managed to present the intricate geometry, which was even too much for Poincaré as his quotation shows, in beautiful pictures which should be accessible to everyone with an interest in science. Applications to biology and medicine may be found in (Glass and Mackey 1988). Chaos theory is very popular also because its geometrical structures (fractals, strange attractors) are of a truly exotic beauty. Particularly remarkable is the combination of beauty and readability in (Briggs 1992). For statistical mechanics and thermodynamics and their philosophical implications, (Lindsay and Margenau 1957) is still unsurpassed.

3.3 Probability

God does not throw dice.

Albert Einstein

Nor is it our business to prescribe
to God how He should run the world.

Niels Bohr

A simple and extremely instructive example of an unstable motion is *throwing a die*. The die is supposed to be a perfect, absolutely homogeneous cube, whose faces are numbered 1, 2, 3, 4, 5, 6.

If we throw it, it will come to rest showing, say, face 3. If we throw it again, trying to repeat the first throw as accurately as possible, it may show a 6 (Fig. 3.4). The *initial conditions* defined by the way of throwing may be almost identical; nevertheless the results will be quite different and practically completely *independent*: instead of a 6, we might as well have got a 4 or a 2.

This is a characteristically instable situation: an arbitrarily small difference of initial conditions will give completely different and independent results. This is the typical situation of a chaotic motion described in sec. 3.2, Fig. 3.4 corresponding fully to Fig. 3.3.

Even if we replace the human hand by a dice-throwing machine, the initial conditions will never be exactly the same, and the result is practically unpredictable. Theoretically its motion is determined by classical mechanics (if also the impact of the air molecules is considered a classical phenomenon), but prediction is hopeless. Laplace's demon,

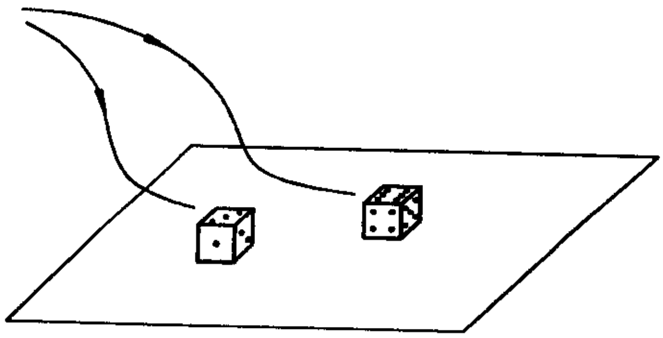

Figure 3.4: Throwing dice

after having worried about the imprecise initial conditions, is then additionally bothered by Maxwell's demon (responsible for air molecules, cf. sec. 4.3).

The result of the fight between the two demons is a completely random distribution of the results of the die: face 1 is as probable as any other face. We may say that all faces have equal *probability*

$$p_1 = p_2 = p_3 = p_4 = p_5 = p_6 = \frac{1}{6} \quad . \tag{3.21}$$

We see that Newton's laws, though theoretically applicable, are practically useless. *Exit* Newton, and *Symmetry* steps in and produces the result (3.21). More prosaically, determinism loses importance and symmetry takes over, producing *order out of chaos.*

(This is the reason why probability is treated here rather than in Chapter 2 where logically it would seem to be better placed.)

Had the die been *loaded*, then, of course, symmetry would have been destroyed and the probability of the various faces would be different. (We, of course, would never use such a dirty trick!)

Such assumptions of equal probability, based on symmetry, were used by Blaise Pascal (1623–1662) and contemporaries for a mathematical theory of games of chance. This was the foundation of the *mathematical theory of probability.* Laplace has perfected this symmetry-based theory.

Here the important concept of *symmetry* appears for the first time. A cube is symmetric because its six faces are geometrically equivalent: they can only be distinguished by marking them with dots, from one to six. If the faces were unmarked, then one could not distinguish a cube lying on face no. 2 from a cube lying on face no. 5. So much about *geometrical* equivalence or symmetry. A cube is also *physically* symmetric if it is made of a ***homogeneous material***: this is what we mean by an unloaded die. A coin is symmetric if we disregard the inscriptions on the two sides: then we could not distinguishing a coin showing "head" from a coin showing "tail": we would in both cases see identical circles. We shall meet symmetry again; see secs. 3.6 and 4.2.

We have said that throwings of various faces were independent events. *Statistical independence* is a basic concept, though it by no means always holds. We shall, however, assume independence unless the contrary is asserted.

Throwing a 3 *or* a 5 has a probability which is the *sum*:

$$p(3 \vee 5) = p_3 + p_5 \quad . \tag{3.22}$$

Throwing a 3 *and* then a 5 is the *product*:

$$p(3 \wedge 5) = p_3 p_5 \quad . \tag{3.23}$$

These formulas do not presuppose equal probabilities (3.21), but they do presuppose independence.

Now we remember symbolic logic, eq. (2.9) on p. 28. The "logical sum" of two propositions was symbolized by "$\vee$", and the "logical product" by "$\wedge$". Now the *probability of a logical sum is the sum of probabilities* (3.22), and *the probability of a logical product is the product of probabilities* (3.23).

Probability 1 corresponds to certainty, and probability 0 to impossibility, and

$$0 \leq p \leq 1 \quad . \tag{3.24}$$

Obviously

$$p_1 + p_2 + p_3 + p_4 + p_5 + p_6 = 1 \quad . \tag{3.25}$$

Thus the probabilities may be considered generalizations of or interpolations between the truth values 0 and 1, cf. (2.15) on p. 41. (Note the conflict of notations: in sec. 2.1, "p" stands for "proposition", here it denotes "probability". As a temporary compromise, we have in (2.15) symbolized probability by "P", but "p" is generally used in probability theory.)

This can also be nicely expressed in the language of set theory (Fig. 3.5). Throw a small particle at random in such a way that it lands on set A with probability $p(A)$ and on set B with probability $p(B)$. Both events may be considered independent if the two sets are

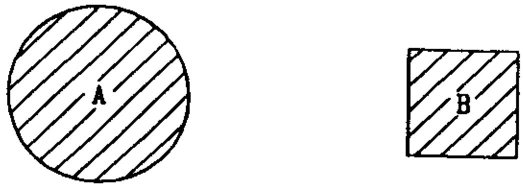

Figure 3.5: Two disjunct sets A and B

disjunct. The union $A \cup B$ consists of both sets A and B taken together. Then

$$p(A \cup B) = p(A) + p(B) \tag{3.26}$$

in analogy to (3.22); cf. (2.12) on p. 30. The corresponding relation

$$p(A \cap B) = p(A)p(B) \quad \text{(wrong!)} \tag{3.27}$$

unfortunately does not hold since the intersection $A \cap B = 0$ for disjunct sets. Here $p(A \cap B)$ would mean the probability that the particle lands *simultaneously* on A and B, which is clearly impossible, so that $A \cap B = 0$ implies $p(A \cap B) = 0$.

Remark on terminology. The terms "probabilistic", "statistic", "stochastic", and "random" have more or less the same meaning and are frequently used interchangeably.

Relative frequencies. Let us take an even simpler example, tossing a coin. For an ideally symmetric coin the probabilities p_1 of head and p_2 of tail are clearly equal:

$$p_1 = p_2 = \frac{1}{2} \quad . \tag{3.28}$$

If we throw the coin, say, a thousand times, there should be roughly 500 heads and 500 tails. In a real coin tossing experiment we may get, say, 484 heads and 516 tails. Thus the *relative frequencies* of heads and tails are

$$\begin{aligned} f_1 &= \frac{484}{1000} = 0.484 \quad , \\ f_2 &= \frac{516}{1000} = 0.516 \quad . \end{aligned} \tag{3.29}$$

If we throw 10000 times, we might get

$$\begin{aligned} f_1 &= \frac{5032}{10000} = 0.5032 \quad , \\ f_2 &= \frac{4968}{10000} = 0.4968 \quad , \end{aligned} \tag{3.30}$$

which is clearly closer to p_1 and p_2. It may be expected that, in some sense, for $n \to \infty$ throws

$$\begin{aligned} \lim_{n\to\infty} f_1 &= p_1 = 0.5 \quad , \\ \lim_{n\to\infty} f_2 &= p_2 = 0.5 \quad . \end{aligned} \tag{3.31}$$

In the so-called *frequency theory* of probability proposed by Richard von Mises since 1928, it was suggested to define probabilities empirically by such a limit

$$p \equiv \lim_{n\to\infty} f \quad . \tag{3.32}$$

This, however, meets with mathematical difficulties because the infinite limit does not obey one of the more common limit definitions of mathematics, and furthermore, it is not possible to perform infinitely many coin tosses or similar procedures.

It is mathematically simpler and more elegant to introduce the concept of probability *axiomatically*. This was done by A.N. Kolmogorov in 1933. Here the probabilities were introduced generally, without specifying their numerical values, but subject to axioms such as (3.22) and (3.23). Only later, approximate numerical values for them are found *a posteriori* as relative frequencies such as (3.30), unless they were not anyway given *a priori* by symmetry considerations (dice, coins).

The mathematical theory of probability has been developed to a high mathematical level, including random functions (stochastic processes) and Hilbert space techniques. Such techniques are, for instance, applied in geodesy to determine the irregular gravitational field of the Earth. This is called *least–squares collocation* and consists in an extension of least–squares adjustment (sec. 2.6) to infinite–dimensional Hilbert space. Only for curious specialists we mention as reference: H. Moritz: "*Advanced Physical Geodesy*", Wichmann, Karlsruhe, 2nd edition, 1990.

Interpretations of probability. In sec. 2.4 we have already briefly introduced *subjective probability*, expressing a degree of reasonable belief, or just a degree of incomplete knowledge or of ignorance. Of such character are the "probabilities of rain" given by American weather forecasts mentioned in sec. 2.4.

The classical Laplace interpretation is clearly intended to be *objective*. When I calculate my chance to gain in gambling to be 95% ($p = 0.95$), then I am not satisfied with this nice abstract result of mathematics: I expect to gain concrete money.

Are physical probabilities subjective concepts or objective features of nature? Consider statistical mechanics. In principle, presupposing the validity of classical mechanics, we could calculate the trajectories of all molecules without needing statistics. Statistics is needed because we cannot do this in practice. Hence we do introduce statistics just because of our inability or ignorance? This would indicate that our probabilities are more or less subjective.

On the other hand, statistical mechanics provides important "*emergent*" concepts such as temperature or entropy, and an elegant theory of thermodynamics has been developed on an axiomatic basis, without needing mechanics or statistics. It seems clear that temperature or entropy are objective "integral" properties of nature, and if they are derived by statistics, this statistics should be more or less "objective" as well. By the way, the derivation of thermodynamics from statistical mechanics is a beautiful example of the emergence of a *macro-law* from a *micro-law*. This is another example of *order out of chaos*.

A hundred years after Boltzmann, these questions are still being discussed. To be sure, the mathematical formalism and its results are completely unaffected by these "philosophical" discussions. Most working physicists could not care less whether their probabilities are subjective or objective. Weizsäcker (1985, p. 100) writes: "The concept of probability is one of the most striking examples for the 'epistemological paradox' that we can apply our basic concepts successfully without really understanding them."

Whether "deterministic chaos" on the basis of classical mechanics "really" introduces an objective probabilistic element into nature, is still an open problem under discussion. Every physicist, however, agrees that *quantum theory* does introduce objective probability into physics: quantum fluctuations form the basic substratum of our world.

Objective probability has been vigorously defended, also in quantum theory, by Sir Karl Popper (Miller 1985, sec. 15). He calls it *propensity* and interprets it in the sense of Aristotle's *potentialities* (possibilities) which are not all realized but are nevertheless *properties of nature*. In the progress of time, potentialities become actualities.

Summary. Probabilities have different interpretations, which are presumably all needed.

(A) *Probabilities of sets.* The current standard mathematical theory of probability, based on Kolmogorov's axiom system, is considering probability as a *measure of sets.* Any system of numbers which satisfies Kolmogorov's axioms is a possible system of probabilities.

Actual *estimations* of probabilities are done in two principal ways:

(1) By *symmetry considerations.* This is easy in the case of dice or coins, but may even be possible in complicated physical applications.

(2) By *relative frequency.* The toss of a coin regards the toss under consideration as one case out of an *ensemble* of 1000 or 10000 tosses, cf. equations (3.29) or (3.30) above. Similarly, in physics, our "real" physical system may be considered one out of a fictitious *ensemble* of possible "similar systems". This is the basis of Gibbs' approach to statistical mechanics (Lindsay and Margenau 1957, sec. 5.5).

Concerning the *physical reality* of probability or statistical considerations, there are two possibilities:

(a) Probability is only a function of *ensembles* of physical systems; probability considerations, such as in statistical mechanics, are only done statistically because a (deterministic) treatment is too complicated for us. Probability is a mathematical tool rather than a physical reality. Hence it may be regarded as *subjective*, at least to a certain extent.

(b) *Propensity*: this type of probability is a *physical property* of a single physical system, as *objective* as its mass, energy, or velocity.

(B) *Probability of propositions.* Mathematically they are very similar to probabilities (A), because the logical calculus of propositions is very similar to the logical theory of sets (sec. 2.1). "Subjective" or *subjectivist* probabilities of Carnap and others are of this type. (The forecast: "There is a 20% probability of rain for tomorrow" is a sentence, or in logical terms, a "proposition".)

Degrees of credibility. Not everything which we call probability must have a numerical value, or must be capable of being expressed numerically. If we say that all our knowledge is only probable, if we believe that the theory of relativity is very probably an outstanding theory, if I say that my train next day will probably run reasonably on schedule, it is difficult if not impossible to assign numerical values to the "probability" expressed by such statements. We instinctively act on beliefs with a high subjective degree of credibility as if they were absolutely true, and we disregard theoretical possibilities which are very small. When I go to work by car I know that I may have an accident. I take this into account in a reasonable way, by insuring my car, having my papers in order, and driving carefully. Having done this, I act as if this eventuality will not occur.

If I kept in mind all the possible events which theoretically might happen, but with a very low probability, then I would "probably" turn crazy or at least become a "professional worrier". This presumably is what Bishop Butler had in mind when he said that *probability is the guide of life* (Russell 1948, Part V, Chapter VI, p. 398).

In real life there is no absolute logical certainty, in the same way as there are no real mass "points", ideal straight lines or ideally exact measurements, cf. sec. 2.4.

Suggested additional reading. Probability, especially of the subjective type, frequently is treated together with induction (to be considered in sec. 3.9). Our standard reference (Lindsay and Margenau 1957) is slightly out of date on this topic but nevertheless worth reading. There are many excellent books on mathematical probability. An easy and delightful brief introduction by the most outstanding Russia specialists is (Gnedenko and Khinchin 1962). Geophysicists will not want to miss (Jeffreys 1961, 1973). A recent excellent discussion of all interpretations and their philosophical aspects is (Cohen 1989). Almost all aspects of probability in their historical development from Blaise Pascal to Niels Bohr are discussed with relatively little mathematics but with beautiful physical intuition in (Ruhla 1992). The remarks in Weizsäcker (1985, Chapter 3; 1992, Chapter 4) are brief but profound.

3.4 The theory of relativity

Henceforth space by itself, and time by itself,
are doomed to fade away into mere shadows,
and only a kind of union of the two
will preserve an independent reality.

Hermann Minkowski

Special relativity

Einstein's special theory of relativity deals with *inertial systems.* An inertial system according to Newton's theory is a system on which