II

Groups and Statistical Mechanics

At about the beginning of the present century, two scientists, one in the United States and one in France, were working along lines which would have seemed to each of them entirely unrelated, if either had had the remotest idea of the existence of the other. In New Haven, Willard Gibbs was developing his new point of view in statistical mechanics. In Paris, Henri Lebesgue was rivalling the fame of his master Emile Borel by the discovery of a revised and more powerful theory of integration for use in the study of trigonometric series. The two discoverers were alike in this, that each was a man of the study rather than of the laboratory, but from this point on, their whole attitudes to science were diametrically opposite.

Gibbs, mathematician though he was, always regarded mathematics as ancillary to physics. Lebesgue was an analyst of the purest type, an able exponent of the extremely exacting modern standards of mathematical rigor, and a writer whose works, as far as I know, do not contain one single example of a problem or a method originating directly from physics. Nevertheless, the work of these two men forms a single whole in which the questions asked by Gibbs find their answers, not in his own work but in the work of Lebesgue.

The key idea of Gibbs is this: in Newton's dynamics, in its original form, we are concerned with an individual system, with given initial velocities and momenta, undergoing changes according to a certain system of forces under the Newtonian laws which link force and acceleration. In the vast majority of practical cases, however, we are far from knowing all the initial velocities and momenta. If we assume a certain initial distribution of the incompletely known positions and momenta of the system, this will determine in a completely Newtonian way the distribution of the momenta and positions for any future time. It will then be possible to make statements about these distributions, and some of these will have the character of assertions that the future system will have certain characteristics with probability one, or certain other characteristics with probability zero.

Probabilities one and zero are notions which include complete certainty and complete impossibility but include much more as well. If I shoot at a target with a bullet of the dimensions of a point, the chance that I hit any specific point on the target will generally be zero, although it is not impossible that I hit it; and indeed, in each specific case I must actually hit some specific point, which is an event of probability zero. Thus an event of probability one, that of my hitting *some* point, may be made up of an assemblage of instances of probability zero.

Nevertheless, one of the processes which is used in the technique of the Gibbsian statistical mechanics, although it is used implicitly, and Gibbs is nowhere clearly aware of it, is the resolution of a complex contingency into an infinite sequence of more special contingencies—a first, a second, a third, and so on—each of which has a known probability; and the expression of the probability of this larger contingency as the sum of the probabilities of the more special contingencies, which form an infinite sequence. Thus we *cannot* sum probabilities in all conceivable cases, to get a probability of the total event—for the sum of any number of zeros is zero—while we *can* sum them if there is a first, a second, a third member, and so on, forming a sequence of contingencies in which every term has a definite position given by a positive integer.

The distinction between these two cases involves rather subtle considerations concerning the nature of sets of instances, and Gibbs, although a very powerful mathematician, was never a very subtle one. Is it possible for a class to be infinite and yet essentially different in multiplicity from another infinite class, such as that of the positive integers? This problem was solved toward the end of the last century by Georg Cantor, and the answer is "Yes." If we consider all the distinct decimal fractions, terminating or nonterminating, lying between 0 and 1, it is known that they cannot be arranged in 1, 2, 3 order—although, strangely enough, all the *terminating* decimal fractions can be so arranged. Thus the distinction demanded by the Gibbs statistical mechanics is not on the face of it an impossible one. The service of Lebesgue to the Gibbs theory is to show that the implicit requirements of statistical mechanics concerning contingencies of probability zero and the addition of the probabilities of contingencies can actually be met, and that the Gibbsian theory does not involve contradictions.

Lebesgue's work, however, was not directly based on the needs of statistical mechanics but on what looks like a very different theory, the theory of trigonometric series. This goes back to the eighteenth-century physics of waves and vibrations, and to the then moot question of the generality of the sets of motions of a linear system which can be synthesized out of the simple vibrations of the system—out of those vibrations, in other words, for which the passing of time simply multiplies the deviations of the system from equilibrium by a quantity, positive or negative, dependent on the time alone and not on position. Thus a single function is expressed as the sum of a series. In these series, the coefficients are expressed as averages of the product of the function to be represented, multiplied by a given weighting function. The whole theory depends on the properties of the average of a series, in terms of the average of an individual term. Notice that the average of a quantity which is 1 over an interval from 0 to A, and 0 from A to 1, is A, and may be regarded as the probability that the random point should lie in the interval from 0 to A if it is known to lie between 0 and 1. In other words, the theory needed for the average of a series is very close to the theory needed for an adequate discussion of probabilities compounded from an infinite sequence of cases. This is the reason why Lebesgue, in solving his own problem, had also solved that of Gibbs.

The particular distributions discussed by Gibbs have themselves a dynamical interpretation. If we consider a certain very general sort of conservative dynamical system, with N degrees of freedom, we find that its position and velocity coordinates may be reduced to a special set of $2N$ coordinates, N of which are called the generalized position coordinates and N the generalized momenta. These determine a $2N$-dimensional space defining a $2N$-dimensional volume; and if we take any region of this space and let the points flow with the course of time, which changes every set of $2N$ coordinates into a new set depending on the elapsed time, the continual change of the boundary of the region does not change its $2N$-dimensional volume. In general, for sets not so simply defined as these regions, the notion of volume generates a system of measure of the type of Lebesgue. In this system of measure, and in the conservative dynamical systems which are transformed in such a way as to keep this measure constant, there is one other numerically valued entity which also remains