

Mathematical Classification of Complete Chaos

Michael Keane

¹ CWI, P. O. Box 94079,
1090 GB Amsterdam, The Netherlands

² Delft University of Technology
2628 CD Delft, The Netherlands

In this essay, we attempt to give a brief survey of the ideas developed over the past years by mathematicians in an attempt to explain the apparent appearance of randomness in deterministic situations. The concepts discussed have a long history, and a detailed investigation would be necessary to give the exact origins, which will not be attempted here. Also no proofs will be given, as there are sufficient expositions at many levels in the literature.

Let us begin by explaining the elementary ideas behind the apparent appearance of randomness. Consider a very simple dynamical system, whose states are the complex numbers of modulus one, and in which the discrete passage of time is given by moving from the state z to the state z^2 . In addition, we assume that the system is in statistical equilibrium, given by the uniform probability distribution on the set of states of the system. This probability distribution is easily seen to be invariant with respect to the passage of time. This simple dynamical system is clearly deterministic, in that complete knowledge of the state of the system at a given time t implies complete knowledge of its state at any time in the future. Nevertheless, there is an element of randomness if complete knowledge of the state is not assumed. Let us suppose that the system is hidden from the observer, who is told at a given time only whether the state z is in the right half or in the left half of the complex plane. Then the observer sees a sequence of zeroes and ones (say, zero for right half and one for left half), and thus progressively gathers information about the beginning state. However, at each observation, the probability of seeing a zero or a one is exactly one-half, independent of the previous observations, so that the observer's knowledge of the future remains completely obscured. The sequence of zeroes and ones which is seen has the same statistical properties as a sequence of random tosses of a fair coin, making the system appear completely random. It is this situation which has been studied mathematically, and we can formalize it as follows.

A *dynamical system* consists of a compact metric space X , together with a homeomorphism S from X to itself and a S -invariant probability measure μ

defined on (the Borel subsets of) the space X . For simplicity we assume that S is invertible – in the above example, this is not satisfied, but this should be considered as a technical detail. Thus by definition, a dynamical system is deterministic, and complete knowledge of the state (a point x of X) at a given time entails knowledge of the state at any time in the future or past, namely, at a given time n the system is in the state $S^n(x)$, if it is in state x at time zero.

Next let us introduce the dynamical systems corresponding to complete randomness, which are called *Bernoulli schemes*. Suppose that $p = (p_1, \dots, p_n)$ is a probability vector, and let Y be the space of all doubly infinite sequences whose elements are taken from the finite alphabet consisting of the symbols 1 to n , n being a fixed positive integer. The left shift operator T acts naturally on Y , by shifting each coordinate of a point one position to the left, and the probability measure ν defined on Y to be the infinite product of the probability p on each coordinate, is S -invariant. Thus we have constructed in a very simple manner a dynamical system in the sense of the above definition, which we recognize as a model for independent repetitions of a single experiment with n outcomes whose corresponding probabilities are given by the vector p .

From the example given at the beginning it should be clear that the dynamical system whose states are complex numbers of modulus one is essentially the same system as a (one-sided) Bernoulli scheme. We now give a mathematical meaning to this notion of similarity. Two dynamical systems are said to be *metrically isomorphic* if there exists a measurable isomorphism between their state spaces which commutes with the corresponding homeomorphisms denoting the passage of time, with probability one. In this definition, the state spaces are considered to be probability spaces, so that *measurable isomorphism* means that there is a probability preserving measurable map (defined almost everywhere), invertible almost everywhere, between them. Thus if two systems are metrically isomorphic, then by relabeling the points of one of the systems (using as labels the points of the other system), the systems become indistinguishable with probability one.

The concept of metric isomorphism has turned out to be very useful in the study of dynamical systems. Only very recently has it been shown by Hungarian mathematicians that the systems consisting of finitely many hard spheres in a box (discretized in time) are metrically isomorphic to Bernoulli schemes. In general, producing metric isomorphisms is a very difficult task, and the constructions available seem to give little insight into the physical nature of the correspondances. The concept is considered by many to be too general to be applicable in real world situations, as it does not take into account the “nearness” of states – a measurable mapping, even if it is measure preserving, has nothing to say about the underlying distances between states. For these reasons, it seems to be desirable to introduce a more stringent notion of identification of systems, which we now discuss.

At first glance, one might expect that the proper concept to require for classification purposes is the continuity of the identification mapping. However, this turns out to be too strong for our purpose, and practically all systems are different if we require continuous isomorphisms. The proper concept seems to be that of *finitary isomorphism*, in which we require that the identification mapping is continuous only after removal of sets of measure zero in each of the state spaces. If we consider again the initial example, then we can see that the identification mapping is indeed not continuous, but only finitary. In fact, if z is a complex number of modulus one, then it can be written as $\exp(2\pi ix)$, with x in the unit interval, and then the sequence of zeroes and ones corresponding to z is just given by the binary expansion of x . The mapping which associates to each x in the unit interval its binary expansion is neither continuous nor invertible, essentially because each dyadic rational number has two different expansions. However, if we remove the dyadic rationals from the unit interval, and the corresponding sequences (which are eventually either all zeroes or all ones) from the sequence space of zeroes and ones, then this mapping becomes continuous, so that here we indeed have a finitary isomorphism. This proves also the utility of the finitary notion, at least by example, since we must admit that binary expansions are useful descriptions of real numbers, even if some continuity is lacking.

Next we would like to discuss the problem of the title. Given the definitions above, it seems clear that the Bernoulli schemes are suitable models for complete chaos. But there are very many Bernoulli schemes, one for each choice of the probability vector p . Let us introduce the notation $BS(p)$ to designate the Bernoulli scheme constructed using the probability vector p . We can then raise the following

QUESTION: When are the Bernoulli schemes $BS(p)$ and $BS(q)$ finitarily isomorphic?

It turns out that this question has a simple answer. To each probability vector p we can associate a number, $h(p)$, called the *entropy* of p , defined by

$$h(p) = - \sum_{i=1}^n p_i \log p_i,$$

and then we can prove that $BS(p)$ and $BS(q)$ are finitarily isomorphic if and only if $h(p) = h(q)$. Intuitively, the entropy, which can more generally be defined for any dynamical system as defined above, measures the amount of information per unit time obtained by observing the system through a finite partition of the state space, so that it becomes an *isomorphism invariant* for dynamical systems. For the systems which are completely random, it is the only number needed to characterize them.

A complete mathematical proof of the above statement is beyond the scope of this exposition, although it is considerably simpler than the proofs of metrical isomorphism and gives more information. However, we can try to convey some

of the flavor of the proof as follows. Suppose that p and q are two probability vectors with the same entropy, and denote by X and Y the sequence spaces of the corresponding Bernoulli schemes $BS(p)$ and $BS(q)$. Suppose also for simplicity that the symbol 1 possesses the same probability in both schemes: $p_1 = q_1$. We want to describe the renaming procedure, which takes a point x in X and turns it into a point y of Y . The first thing to do is to look at the infinite sequence x and copy all occurrences of the symbol 1 to the corresponding places in the point y we are looking for. With probability one, this now divides the remaining parts of x into finite strings of the remaining symbols. Using concepts of information theory, it is then possible to provide a sequential procedure to successively copy larger and larger blocks of these finite strings to the corresponding places in y , taking care that shifted objects are copied in the same manner.

Finally, we remark that the notion of finitary isomorphism is closely connected with the classes of observables which we consider. For this purpose, an *observable* should be thought of as a function defined on the state space of the dynamical system. In metric isomorphisms, the natural class of observables is the set of measurable functions, whereas in continuous isomorphisms the natural class of observables would be the continuous functions. For finitary isomorphisms, the natural class of observables is the set of functions which are continuous after removal of a set of measure zero, allowing many indicator functions of sets (those which have boundary measure zero) but still preserving the physical nature of closeness with probability one. Under such an isomorphism, these “almost continuous” functions are in one-to-one correspondence and hence their statistics behave in the same way in isomorphic dynamical systems. This property emphasizes the physical nature of the underlying systems.