数理解析研究所講究録
1011巻 1997年 42-57

Randomness Revisited

Physics and Beckman Institute, UIUC[1]
Y. Oono

I Introduction–Characteristics of Nonlinear Systems

Nonlinear systems are systems that are not linear. Hence, a 'nonlinear system' means that a system whose particular description [1] does not obey the superposition principle. 'A system is nonlinear' is a shorthand statement of this.

The word '*non*linear' may not be, however, a fully satisfactory characterization of 'nonlinear' systems, because the word does not positively characterize nonlinear systems; it characterizes a system by the absence of linearity, just as infinity is characterized by the lack of finiteness. In this case Dedekind supplied a positive definition of infinity with the aid of one-to-one correspondence between a set and its proper subset. Is there any positive characterization of nonlinear systems?

Why does the superposition fail in 'non-linear' systems? Because we cannot separate different scales (length, time, etc.). Consider a nonlinear partial differential equation such as the time-dependent Ginzburg-Landau equation. Decomposition into different Fourier modes is useless to solve it because of 'mode-coupling.' We cannot disentangle small and large scales in contrast to the case of, say, the linear diffusion equation. This suggests a positive definition of nonlinear systems as <u>scale-interfering systems</u> [2]. Scale interference causes nonlinear systems to exhibit various nontrivial and often unexpected phenomena such as phase transition, critical

[1]University of Illinois at Urbana-Campaign

phenomenon, chaos, etc. 'Nontriviality' of nonlinear systems is solely due to this scale interference that promotes the unknowable relevant to what we can observe at our (space-time) scale. The effect of noise illustrates the point in a rather trivial fashion; it is obvious that noise has no effect at all on the expectation values of the quantities governed by linear evolution equations, but often has marked effects on nonlinear systems as exemplified by critical phenomena.

In this report, I wish to consider a foundational aspect of randomness. Randomness becomes much more important to nonlinear systems than to linear systems because of scale interference. Hence, much more serious attention should be paid to the fundamental nature of randomness in Nature. As the reader can easily infer, the scale interference makes the world we observe not-self-contained (open to the unknowable universe). In the next section, the real significance of chaos is discussed from this point of view. In Section III, a fundamental nature of noise and randomness to natural science is stressed. In Section IV characterization of randomness by van Lambalgen is reviewed (with a theoretical physicist's gross distortion), and its possible serious consequences are discussed. Needless to say, the problem I wish to consider is so fundamental that I will not be able to say any positive comment. Rather, I wish to point out that fundamental scientists (not mathematicians) should pay due attention to, e.g., the choice of axiomatic systems of set theory as a respectable empirical question just as the question whether the world is Euclidean or not has been regarded so.

Before going into the main topic, it may be useful to recall the history of modern mathematics. Most people agree that Fourier's proposal to use Fourier series to solve the diffusion equation made an epoch. Cauchy was spurred to put analysis on a firmer ground; Riemann had to invent his integral to compute Fourier coefficients according to Fourier's prescription. Along this line lies the fundamental work of

Lebesgue. On the other hand, the question of uniqueness of Fourier expansion drove Cantor to think about the distinction between denumerable and non-denumerable sets. To secure his theory of transfinite numbers, he had to invent set theory. To secure his well-ordering proof, Zermelo had to formalize axiomatic set theory with the axiom of choice (AC). We see that the (foundation of) mathematics we currently use has largely been molded to solve various basic problems of Fourier expansion and transformation. In other words, linearity molded our mathematics. It may be worth reevaluating this fact before going into the study of nonlinear systems.

II Chaos and Randomness

Perhaps the most dramatic effect of scale interference is 'chaos.' Intuitively, 'chaos' is a 'random' behavior of a deterministic dynamical system. There have been many characterizations of chaos. Here I choose one of the oldest definitions [3], because it is naturally related to algorithmic randomness [4] (and also almost no one has understood the significance of the definition; see the remark below).

If one wishes to proceed as intuitively as possible temporarily avoiding formalization of randomness, it is a sensible idea to use an intuitively clearly 'random' process such as the coin-tossing process of a fair coin (the one-sided $B(1/2, 1/2)$ process or the full shift on $\{0,1\}^{\boldsymbol{N}}$) to formalize 'chaos.' Roughly speaking, if we can find a reasonable correspondence between the sample histories of the coin-tossing process and (a subset of) trajectories of a given deterministic dynamical system, the dynamical system may be said to exhibit chaos [5].

Definition [chaos]. Let f be an endomorphism of a set (usually a region) Γ into itself. We say that the map f exhibits *chaos*, if there exists $n \in \boldsymbol{N}$ and $A \subset \Gamma$ such that f^n restricted to A is isomorphic to the shift dynamical system σ on $\{0,1\}^{\boldsymbol{N}}$.

That is, the following diagram is commutative [6]:

$$\begin{array}{ccc} A & \xrightarrow{f^n} & A \\ \downarrow{\scriptstyle\varphi} & & \downarrow{\scriptstyle\varphi} \\ \{0,1\}^{\boldsymbol{N}} & \xrightarrow{\sigma} & \{0,1\}^{\boldsymbol{N}} \end{array}$$

The following theorem [8] summarizes nice features of the definition [9]:

Theorem. *Let $F : I \to I$ be a C^0-endomorphism of an interval I. Then, the following (1) to (4) are equivalent [12]*

(1) F exhibits chaos.

(2) F has a periodic orbit of period not equal to an integer power of 2.

(3) There exists a positive integer m such that F^m has a mixing invariant measure.

(4) F has an invariant measure with a positive Kolmogorov-Sinai entropy. □

Notice that Ornstein and Weiss [13] identify chaos and positivity of entropy.

Remark. The equivalence of (1) and (2) is the well-known statement: *Period* $\neq 2^n$ *implies chaos* [3]. This may sound a straightforward generalization of the famous theorem due to Li and Yorke [14]: *Period three implies chaos*, but there is an important distinction that is usually totally ignored [15]. That is, the definitions of chaos in the above theorem and in the Li-Yorke theorem are distinct. Chaos implies the Li-Yorke chaos, but the converse is untrue, as can be seen easily from the equivalence to (4). Obviously, the above theorem is *not* a corollary to the Li-Yorke theorem (even with the aid of Šarkovski's theorem [18]). There is also a grave defect in the definition of the Li-Yorke chaos. The scrambled set carrying the Li-Yorke chaos is often measure zero [19], so that the Li-Yorke chaos is NOT observable. In contrast, the set A is with positive measure whenever there is an absolutely continuous invariant measure. The Li-Yorke chaos never pays attention to randomness, which causes its defect. □

The above characterization may be stated that if the trajectories have natural relation to algorithmic random sequences, then the system is chaotic. This obvious restatement has a strong support of Brudno's theorem [20]. This theorem implies that the loss of initial information (measured by the Kolmogorov-Sinai entropy) and impossibility of its short coding (measured by the Kolmogorov complexity) are equivalent. Since the initial information is constantly being lost in chaotic systems, we must provide infinitely precise initial information to specify a trajectory. Therefore, no information-compression is possible, and no prediction of the long future is possible.

Incidentally, Brudno's theorem has some practical relevance, although I have never seen the theorem quoted in physics literature except in a recent one [21] (except ours [22]). Suppose we wish to compute a trajectory of a dynamical system whose time one map is given by T. To code a trajectory starting from a point x for a period t, we need roughly $tK(x,T)$ of symbols, where $K(x,T)$ is the Kolmogorov complexity of the trajectory starting from x. $tK(x,T)$ may be interpreted as a computational cost of the trajectory starting ¿from x for time t. That is, $K(x,T)$ may be understood as the necessary cost to compute one time unit into the future from the present data. This idea has been successfully applied to the Sinai billiard system [23].

III Significance of Randomness

As explained in the previous section, randomness is crucial to chaos. Where does its randomness come from? It is from the initial condition, especially, from its details. Because extremely small scales can interfere our scale (due to generally expansive nature of chaotic dynamics), chaotic systems look random. Notice that such small scales are unknowable. This is not due to our technological limitations but due to absolute limitation. Chaos is conceptually important, because it reminds us of the existence of the unknowable; Chaos magnifies what we can never observe, and because of this the unknowable can have significant effects on what we observe or experience.

To recognize the possibility of analyzing a set of trajectories of a mechanical system into the equation of motion and initial conditions was a major ingredient of the Newtonian revolution. To understand the equation of motion gives us a feeling of understanding the system if it is simple, so that (perhaps reinforced by the elementary mechanics education) we thought the equation of motion was everything. Then, came chaos that exhibits counterexamples to our happy feeling of understanding in terms of the equation of motion. In order to specify a state of a chaotic system we must know much more than its equation of motion. One can say a chaotic system is a system for which the information in the initial condition dominates the required information to specify its behavior. Notice that the parallelism between chaos and algorithmic random sequence is really good.

The significance of noise = the unknowable has become easy to grasp even for uncritical thinkers thanks to simple chaotic system examples. This is the most significant implication of chaos. This is, however, not a fundamentally important progress at all, because science is not a democratic business. It is too shallow a view point to advocate that with chaos new era of science dawns.

¿From the application point of view, the significance of chaos is that it is a modifier of randomness given by Nature. The randomness in chaos comes from its initial conditions. One could use neatly modified noise by chaos (like chaotic itinerancy [24]) to accelerate optimization processes. The problem of using chaos as a random number generator is how to minimize the distortion due to deterministic modulation.

In physics, there are other cases where randomness is significant. Everyone is taught the principle of equal probability as a foundational principle of statistical mechanics. Although there have been many people attempting to reduce it to a theorem of mechanics, the attempt is intrinsically meaningless, because the principle is a principle about the initial condition of the mechanical (quantum or classical) system. Every sensible person knows that the initial condition is beyond the control of mechanics. Hence, the principle, which specifies what we mean by random sampling in experiments, is a really fundamental principle, perhaps more so than the law of mechanics.

Quantum mechanics also requires deep understanding of randomness. If the wave contraction due to observation could be explained, for example, by the irreversibility in the observation process, then again we are not free from the nature of initial conditions before the observation.

IV van Lambalgen's Axioms

We must conclude that randomness is a fundamental concept in natural science. In the above the Solomonov-Kolmogorov-Chaitin algorithmic characterization of randomness was adopted. However, this characterization relies on the concept of computation, or the Church thesis [25].

Is the concept of computation or algorithm more fundamental than randomness? There might be a viewpoint that computation or the Turing machine is a device to formulate the concept of randomness, so the aforementioned question is misplaced. Let us consider the concept of continuity. The concept seems very fundamental, but we now know that topology is a much more fundamental and simpler concept than continuity, so that we never regard topological space as a device to discuss continuity. As can be seen in the definition of the Turing machine or the statement of Church's thesis, computation is not a simple concept (perhaps even ugly).

Also, one must not ignore that there is no guarantee that all the regularities are 'computable regularities' detectable by a universal Turing machine; some inspiration (oracle) might tell us the existence of a different kind of regularity in a given sequence.

This is not an outrageous statement, if mathematical intelligence is, as supposed by Gödel, nonfinitary [26]. Furthermore, 'random sampling' may be done by Nature Herself. In that case, why do we have to assume that Her capability is restricted to computation? I do not adopt a rather extreme view that Nature Herself is a computer, so I feel it unnatural to regard algorithm as a basic concept.

If we wish to regard randomness as fundamental, unless we can find something intuitively more fundamental, we cannot define the concept 'randomness.' Hence, only possible way to mathematize the concept is axiomatization. A very serious attempt the author is aware of is that due to van Lambalgen [27].

The following is a rather degenerate version of his axiomatic effort interpreted in a theoretical physicist's way. Because random sequence is defined algorithmically by its incompressibility, and because our intuitive reason to feel a sequence to be random is that we cannot find any regularity in it, we should axiomatize the concept of 'incompressibility.'

Van Lambalgen introduces the relation R such that $R(x;y)$ may be interpreted as 'x cannot be information-compressed even with the extra information y (or 'oracle y')'. Here, x is interpreted as a 01 sequence (or some symbol sequence) and y may be a set of symbol sequences. Notice that x can be information-compressed if we already know x itself (we have only to say 'the same'), so $R(x;x)$ is untrue. This relation R is specified by the following axioms, which are put very informally here [28]:

R1. There is a sequence that cannot be information-compressed without any 'external information (or oracle). [i.e., $\exists x R(x,\emptyset)$.] ($R(x,\emptyset)$ is henceforth written as $R(x)$.)
This asserts that there is a random sequence. $R(x)$ can be interpreted as the statement that x is random.

R2. If x cannot be information-compressed with the information of y and z, then it cannot be done so with the information of z only. [i.e., $R(x;yz) \Rightarrow R(x;z)$.] (Here y and z may be understood as sets, and yz their joint set, so the order of y and z does not matter.)
This asserts that there is a hierarchy in randomness, although $R(x;zy)$ does not necessarily imply that x is more random than x' such that $R(x';z)$.
R3. If x cannot be information-compressed by the information of y, then x and y are different. [i.e., $R(x;y) \Rightarrow x \neq y$.]

These are very natural axioms.

R4. If a certain relation ϕ of y to x does not allow one to information-compress x, then there is w such that the ϕ-relation of y to w plus an extra information z does not information-compress w. [i.e., $\exists x(R(x,y) \bigwedge \phi(x,y)) \Rightarrow \exists w(R(w,zy) \bigwedge \phi(w,y))$. Here ϕ should not have any parameters other than listed in y.] (In a certain sense, x satisfying $R(x,y)$ is 'more random (lawless)' than w satisfying $R(w)$, because **R2** implies that x satisfies $R(x)$.
Consequently, this axiom demands that there always exists a 'more' random sequence than a given one; Indeed, **R4** implies, with no relation chosen as ϕ, $\exists x R(x;y) \Rightarrow \exists w R(w;zy)$, or, without y $\exists x R(x) \Rightarrow \exists w R(w;z)$ for any z. **R4** may sound artificial. However, if we negate this, then there is some set of sequences A such that there is no x such that $R(x;A)$. That is, if we know A, then any sequence x can be information compressible. Hence, the axiom **R4** implies the 'non-existence of Almighty.'

There are more axioms, but I have mentioned only I need below.

V Consequences of Axiomatization

As briefly explained in the preceding section, the axioms of R are intuitively understandable (appealing) natural requirements. The consistency of these axioms with the usual Zermelo-Fraenkel set theoretical axioms (without AC) (ZFR) is established by van Lambalgen [27, 29].

However, the ZFR is contradictory to AC:

Theorem [van Lambalgen]. There is no choice function on the power set of reals. □

[Outline of proof]. Take an arbitrary sequence y. There is a sequence x which cannot be compressed with the knowledge of y. Let g be a choice function on the power set of reals (this is expressed as a set). Even if we know g, there must be an incompressible sequence x, so that $\{x \mid R(x;g)\}$ is a nonempty set. Hence the choice function g must give an element of this set y: $y = g(\{x \mid R(x;g)\})$. By definition, this implies $R(y;g)$. Intuitively speaking, this is strange, because y is determined by g, so y should be compressible with the knowledge of g. Hence, we arrive at a contradiction. Formally, we can proceed as

$$\exists y R(y;g) \bigwedge y = g(\{x \mid R(x;g)\}). \tag{5.1}$$

With the aid of **R4** (setting $z = g(\{x \mid R(x;g)\})$)

$$\exists y R(y, g(\{x \mid R(x;g)\})g) \bigwedge y = g(\{x \mid R(x;g)\}). \tag{5.2}$$

However, **R2** allows us to replace $yR(y, g(\{x \mid R(x;g)\})g)$ with $yR(y, g(\{x \mid R(x;g)\}))$ in this formula, so that this implies $y \neq g(\{x \mid R(x;g)\})$, a contradiction.

Physics, or more generally, natural science may never require general AC asserting that any transcendental family has a choice function. The assertion may be interpreted as the credo of 'super-realism.' If we admit only the countable axiom of choice (CAC), we could keep all the theorems of analysis needed for physicists, and, simultaneously, can get rid of non-measurable sets or Hausdorff-Banach-Tarski type 'unpleasant' theorems. Hence, physicists may even be delighted with the fall of the full strength AC. However, if CAC fails, we would be in big trouble.

Let $\mathcal{U}$ be a countable family of sets and g be a choice function on it. Let $B = \{x \,|\, R(x, g)\}$. Now, make a countable family of sets joining $\mathcal{U}$ and B, and h be a choice function on it. Of course, $R(h(B); g)$, because $h(B) \in \{x \,|\, R(x, g)\}$. In other words,

$$\exists y R(y, g) \bigwedge y = h(\{x \,|\, R(x, g)\}). \tag{5.3}$$

Hence, with the same logic employed in the proof above,

$$\exists y R(y, h(\{x \,|\, R(x, g)\}) g) \bigwedge y = h(\{x \,|\, R(x, g)\}), \tag{5.4}$$

which leads us to a similar contradiction. Hence, CAC also fails.

VI What can we conclude?

It is hard to conclude or propose something mathematically constructive. In this concluding section, however, I try to extract some lessons.

It seems that CAC and a natural axiomatic system of randomness are not compatible. This is more than we physicists can swallow. If we accept that the hierarchy of randomness has an upper bound (so that we have an almighty oracle set), then certainly we can avoid this 'unpleasant' conclusion. However, to avoid Almighty is a prerequisite of any serious science, so this boundedness by something or someone is unacceptable.

Of course, the reader would say that to accept randomness as fundamental is the source of problem. However, for empirical scientists I believe randomness is more fundamental than some esoteric set theory axioms like large cardinal axioms. If one wishes to stick to the idea that randomness is fundamental, then what are the possible directions?

Van Lambalgen himself discussed various possibilities [27]. He considered the possibility of replacing the extensionality axiom with a weaker version. Under this condition the 'set' B above is not a set, so the demonstrations in the previous section do not work. Besides, AC without extensionality is probably not powerful enough. Of course, there must be a lot of room to modify the simplest axioms for 'randomness.' Perhaps, the current attempt assumes the transcendency of randomness excessively. We could add many fine prints, but I do not wish to modify the simplicity of the above mentioned axiomatic system for R. In any case all the proposals to try to avoid contradiction between the current axiomatic system (ZFC) and randomness seem to suggest modifications of extreme atomistic nature of the current mathematical foundation.

There is a drastic position to deny all the infinity: since we are made of atoms (or elementary particles), and the total number should be finite, because the universe seems to be finite (or in any case we discuss the phenomena in front of us, so we have only to think about bounded domains). Perhaps the claim is true and there is no infinity in Nature. However, to understand is to idealize. Without limiting procedures, it is impossible to idealize. If no infinity is allowed, even thermodynamic phases cannot be defined unambiguously, because we cannot take the thermodynamic limit. Hence, I totally reject any position denying idealization.

To consider what axiomatic system of geometry is adapted to the actual space has been a respectable empirical question. To ask what axiomatic system of the set theory is the most suitable to understand Nature can also be a respectable question. Of course, it has been a respectable question as can be seen in [30] among 'fundamentalists,' but not yet among empirical scientists.

One intriguing point I wish to stress here is that the conceptual problem of randomness was never serious in the study of linear systems. Note (1) ZFC is almost a direct consequence of the quest for the foundation of Fourier expansion theory (i.e., a tool for linear systems), and (2) randomness is never important in linear systems. However, everyone knows that linearity is only due to approximation (except perhaps for the Schrödinger equation), so (1) scale interference implies that the linear expansion tools are no more powerful, and (2) the effect of randomness becomes even qualitatively significant. There might be a more natural foundation of mathematics which can be a suitable tool to understand Nature full of nonlinearity. Perhaps, the current mathematics is artificially constrained by simple minded idealization (called linearization).

Acknowledgements. I acknowledge fruitful discussions with Shin-ichi Sasa. The work is, in part, supported by the National Science Foundation Grant DMR 93-14938.

Bibliography

[1] Just as seen in the so-called Koopmanism, a nonlinear dynamical system can be described by a linear equation such as the Liouville equation, so that there is not an intrinsic linear or nonlinear system. Even if we believe that everything can be described by the Schrödinger equation, we know there are many nonlinear phenomena such as life.

[2] I ignore integrable systems, because they are pseudo-nonlinear systems or disguised linear systems.

[3] Y. Oono, Prog. Theor. Phys. **59**, 1029 (1978).

[4] See, e.g., M. Li and P. Vitányi, *An Introduction to Kolmogorov Complexity and Its Applications* (Springer, 1993) as a convenient reference.

[5] Continuos dynamical systems can be discretized by one of many standard methods, so we need not consider them separately.

[6] Alekseev [7] defined *quasirandom* dynamical systems in terms of a Markov chain with positive Kolmogorov-Sinai entropy. The idea is, in retrospect, the same.

[7] V. M. Alekseev, Math. USSR Sbornik, **5**(1), 73-128 (1968).

[8] Y. Oono and M. Osikawa, Prog. Theor. Phys. **64**, 54 (1980); M. Osikawa and Y. Oono, Publ. RIMS, **17**, 165 (1981).

[9] See also [10]. Block and Coppel [11] define a 'turbulent map' f by the condition that there are two intervals J and K sharing at most one point such that $J \cup K \subset f(J) \cap f(K)$, and say that when f^n is turbulent for some positive integer n, f is chaotic. This usage is exactly the one proposed in our definition. The existence of such J and K is a lemma used in the proof of this theorem.

[10] L. Block, Proc. Amer. Math. Soc., **67**, 357 (1978); ibid., **72**, 576 (197 8).

[11] L. Block and W. A. Coppel, *Dynamics in One Dimension*, Lecture Notes in Mathematics **1513** (Springer, 1992).

[12] (1) $\Rightarrow$ (2), (3), (4) are trivial.

[13] D. Ornstein and B. Weiss, Bull. Amer. Math. Soc. **24**, 11 (1991).

[14] T.-Y. Li and J. A. Yorke, Am. Math. Month. **82**, 985 (1975).

[15] For example, Hao [16] says on p24, "Unaware of the work of Šarkovskii, many equivalent formulations or corollaries of this theorem were suggested in the intervening years, among which the Li-Yorke theorem "Period 3 implies chaos" and the Oono's statement "Period $\neq 2^n$ implies chaos", etc." A similar misunderstanding can be found in Li [17].

[16] B.-L. Hao, *Chaos* (World Scientific, 1984).

[17] S. Li, Trans. Amer. Math. Soc. **339**, 243 (1993).

[18] A. N. Šarkovskii, Ukr. Mat. Z. **16**, 61 (1964); P. Štefan, Commun. Math. Phys. **54**, 237 (1977).

[19] Y. Baba, I. Kubo and Y. Takahashi, Nonlinear Analysis **26**, 1611 (1996).

[20] A. A. Brudno, Trans. Mosc. Math. Soc. **44**, 127 (1983).

[21] A. Crisanti, M. Falcioni, G. Mantica, and A. Vulpiani, Phys. Rev. E **50**, 1959 (1994).

[22] For example, Y. Oono and C. Yeung, J. Stat. Phys. **48**, 593 (1987).

[23] P. R. Baldwin, J. Phys. A **24**, L941 (1991).

[24] I. Tsuda, World Futures **32**, 167 (1991).

[25] R. I. Soare, *Recursively Enumerable Sets and Degrees*, (Springer, 1982).

[26] R. Gandy, "The Confluence of Ideas in 1936," in *The Universal Turing Machine, a half-century survey* (Oxford UP, 1988).

[27] M. van Lambalgen, J. Symbolic Logic **55**, 1143 (1990); J. Symbolic Logic, **57**, 1274 (1992).

[28] There are slightly different versions of axioms given in van Lambalgen's papers. Also here the exposition is informal, so the numbering of the axioms is different from the original versions.

[29] There must be details such as we use the first order logic, etc., but in this informal exposition, I omit them.

[30] P. Maddy, *Realism in Mathematics* (Oxford, 1990).