hardly demonstrate impossibility. In all probability we are unprepared to think realistically about these problems as yet, to distinguish between the sort of differentiation and structure that emerges merely as a function of the power of recursive instructions, and the sort which argues for the sequential masking of old instructions and the uncovering of new instructions. Ulam has described aperiodic crystal-like or dendritic structures of considerable apparent complexity, but generated by the iteration of very simple instructions which recursively copy simple initial configurations.[10]

There is, of course, every reason to suppose that the power of different instructions involved in morphogenesis varies over a wide range. This is, plausibly, the conclusion to which evolutionary arguments would also lead. At one extreme, with very many "embryological" instructions of relatively low power, the set of accessible genetic variations is very large and the average differences among them are small. This would imply the possibility of refined adaptation, copying of the environment, over a narrow range of ecological loci. On the other hand if the number of instructions is small and their power high, the set of accessible variations is small but phenotypic differences will average very large, too large to allow local adaptation or, even, viability in any locus. An obvious condition of optimality would seem, therefore, to be a wide spectrum of powers. The argument is similar to that which one invents for optimal random search patterns, which requires both occasional long leaps and frequent short exploratory steps.

If this argument is correct it implies that one might do well to give a high priority (in the game of research) to models of morphogenesis which require relatively few genetic instructions. And there may be an important point in the study of naive models of the kind referred to above, simple studies of growth and form in gamelike contexts. These will not give much, perhaps, that could sententiously be called "mathematical biology," but they might help us learn to think better about a range of problems that have, until recently, been almost beyond formulation.

---

[10] S. Ulam, "On Some Mathematical Problems Connected with Patterns of Growth of Figures," in *Proceedings of Symposia in Applied Mathematics* (American Mathematical Society, Providence, Rhode Island, 1962), Vol. 14. pp. 215–224.

A Thermodynamic Characterization of Self-Reproduction

P. MORRISON

Department of Physics and Newman Laboratory of Nuclear Studies, Cornell University, Ithaca, New York

> Horses, elephants, and iron,
> Water, woman, man,
> Sticks and stones and clothes are built
> On a different plan.
> —*The Panchatantra**

0. TOWARD A DEFINITION OF LIFE

A wise biologist has reminded us that definition is "as a matter of fact, an excellent heuristic method." Though the mathematician's need for a definition *a priori* has little relevance to the real situation of the sciences of life, it is hard to dispute Lwoff's point.[1] The casting of a definition of life, while it by no means is a necessity for biology, "compels critical consideration of all the terms or aspects of a problem." Three sorts of approaches to the question are current: one which fixes its attention upon holistic aspects,[2] a purely chemical one, which seems less than successful,[2] and that of the microbial geneticists,[1,3] who describe life as an independent, specific, self-replicating system, possessing some sort of continuity with the general organic cycle. (The acid, but perhaps not captious, critics who have observed that mules and Prime Ministers would then generally fail to be classed as alive will not be fully content with the present approach either!) It is the aim of this essay to present a parametric measure of certain quasi-thermodynamic properties of all systems which may be called self-reproducing. The scope of the meaning of self-replication will be wide, so wide that many nonliving systems will clearly allow a measurement

---

* Translated from Sanskrit by Arthur W. Ryder.

[1] A. Lwoff, *Biological Order* (M.I.T. Press, Cambridge, Massachusetts, 1962), p. 1 *ff*.

[2] K. Smith, Am. Scientist **46**, 413 (1958) and references there cited.

[3] S. Luria, *General Virology* (John Wiley & Sons, Inc., New York, 1953).

of the relevant parameters. The various systems present a monotone sequence, running from low values of the parameter up to the maximum of its range, in which the place of the clearly nonliving examples is at the low end, but that of the obviously living examples at the high. A cut of more or less arbitrary position in the sequence would thus part the quick from the dead. Insofar as the proposal is successful, it is to be regarded as an attempt to make objective and quantitative the ideas behind the geneticist's definition referred to above.

Placing self-reproduction so important in the categorization is of course at once to ignore the counterexample of the mule. It ought to be clear, though, that at least the *production* of the mule, or of the zygote past reproductive age, was in itself a self-replicating process of the expected kind. Formally, then, to describe the living as that which *originates* in such a process would escape the contradiction. It should be clear that any effort at structural rather than functional characterization is bound to meet with at least apparent counterexamples. A clock might be precisely described, except for a dust fragment which prevented its going. Its failure, the analog of death, occurs even though the internal structure is all but perfect. Surely no physical description beyond a merely circular attribution of correct function can class a clock stopped by a single speck of dust with a box of unassembled parts, or with a half-kilo brass ingot. There is a way out; the systems described must be regarded, not as unique individuals, but statistically. In the configurational neighborhood of that structure which is a clock stopped by one dust speck lie many other, similar, but not identical clocks —those without a dust speck, or with the speck placed anywhere near, but not at, a crucial interface —which function perfectly well. The familiar experience of striking a sharp blow against the case of a defective mechanism, clock or radio, is a homely demonstration (in time, rather than in ensemble average) of this point. Often clinical death can be reversed by a pacemaker, or a heart massage, or perhaps by even less specific therapy; an organism just dead through mere heart failure is not at all the dead creature after internal lysis has begun. Since it is to be expected that thermodynamic or statistical-mechanical arguments will always contain ambiguities which approach zero only when truly large numbers of systems are considered, we shall rest content with a discussion meant in this statistical sense, and admit freely that individual counterexamples, like nonreproducing but vigorous organisms, or those all but intact and yet plainly "dead," can be cited against our defining parameter. The larger view, which takes into account the whole class of organisms, in its *production* as well as in its *reproduction*, and as well organisms configurationally close, thus perhaps reversing "death," or ending a relatively minor barrier to interspecies fertility, is here implied. This procedure is analogous to the familiar coarse graining of Gibbs. In this coarse-grained sense the parametric sequence will be presented as at least a sufficient criterion of a continuous transition from the nonliving to the patently alive.

1. THE THERMODYNAMICS OF INFORMATION: SEMIPERMEABLE MEMBRANES

For a good many years it has been widely, if not unanimously, recognized that information embodied in any physical form whatever could provide a source of free energy, and hence available work. The truly classical example is of course the separation of mixed gases, which leads to the famous Gibbs paradox. In his illuminating review, Willkie[4] has made clear how the ordinary relation for the free energy change upon chemical reaction

$$a\mathrm{A} + b\mathrm{B} \rightarrow c\mathrm{C} + d\mathrm{D} :$$
$$\Delta F = \Delta F_0 + RT \ln [\mathrm{C}]^c[\mathrm{D}]^d/[\mathrm{A}]^a[\mathrm{B}]^b$$

is to be regarded as the sum of two terms, ΔF_0 the standard free energy change, which expresses the internal energetic and configurational chemical differences between products and reactants at a conventional concentration; and the concentration-dependent logarithmic term, which takes exactly the form of a mixing free energy. Even if the first term vanishes, the existence of any physical differences whatever which enable a distinction to be made—and information gained—between product and reactant molecules leads to a possible yield of work. This requires a concentration change carried on in the typical van't Hoff reaction box fitted with pistons bearing appropriate semipermeable membranes. The dependence on concentration is informational in character, and has nothing whatever to do with the concentration-dependent deviations from ideal solution behavior arising from ionic or similar interaction.

If one takes a simple case of identity change, say the reaction

$$\mathrm{A} + \mathrm{BC} \rightleftarrows \mathrm{AC} + \mathrm{B} ,$$

where the "atom" designated by C is merely a marker containing no appreciable energy of binding, mass, or constraint on configuration, except for its

---

[4] D. R. Willkie, Progr. Biophys. Biophys. Chem. **10**, 260 (1960).

mere presence, the free energy F released per mole of reactant A consumed in coming to the equilibrium concentration, $n_{\mathrm{A}} = n_{\mathrm{B}} = \frac{1}{2}$, is given by

$$F = 2\,RT[(n_{\mathrm{B}} - 1)\ln(1 - n_{\mathrm{B}}) - n_{\mathrm{B}}\ln n_{\mathrm{B}}]_0^{n_{\mathrm{B}}}$$

which is here $2\,RT \ln 2$, or $RT \ln 2$ per reactant. The ability to distinguish the marker position has yielded free energy, but only if semipermeable membranes capable of discriminating all the participant molecules are employed. An experimenter not sufficiently discriminating would be unable to detect the free energy potentially available; it would be made up fully by the undetected mixing entropy of a mixture he was unable to separate. But it cannot be held that this free energy contribution is not real; it is a source of work which is available only to changes carried on under appropriate constraints, but indeed so is every transfer of energy into work.

It is of value to us to look more closely at the problem of realizing the semipermeable membrane in general, for we shall demand very specific membranes. The general thermodynamical theory clearly implies their existence in principle for any difference, no matter how refined. It ought to be possible to describe their construction. In the classical case at least, such membranes must of course be entirely passive. They can be built only out of prescribed potentials, exactly as a confining wall or piston is built out of a more or less rigid repulsive potential hill, entering the molecular Hamiltonian at the prescribed positions in space. It is easy to make a beginning: a potential hill with apertures of chosen size will make a piston permeable to rigid spherical molecules whose diameter is smaller than the aperture. If we relax the rigidity requirement, we can specify the separation factor as a function of temperature. This is no defect in the general utility of the scheme. If we consider molecules of any well-defined, time-independent shape, the same external collision barriers will work, though the aperture may require a complex template, and the time of passage will lengthen. So to discriminate between a pair of molecules of any two shapes seems possible; some dimension of the envelope of the molecule will make a *go no-go* choice. To filter out a thin rod from a shorter one of the same diameter, for example, would require an aperture cut in the wall of a cylindrical cavity which would admit diffusing molecules only with their axes parallel to the cylinder axis. To pass a rod of given diameter, while retaining rods of equal length, but larger diameter, clearly demands, on the other hand, an aperture at the end of a cylindrical cavity. There seems to be no limit to the complexity of the discrimination required as between two alternatives. We will accept as at least plausible the construction of membranes capable of choosing between any two distinct shapes of molecule. Even if the external form of the molecules is identical, their inner components in principle are not entirely without coupling energy to external potentials, so we can imagine a membrane capable of cutting out filled boxes but admitting identical but empty ones; the potential hill is slightly higher when the internal components are present. Whether such a model is realistic for the typically electrical interactions among real molecules is not here under discussion.

But mere discrimination is not enough. We need membranes not only to pass B and retain A, say, but to pass A and retain B. Otherwise the reaction box will not extract free energy. Clearly we cannot expect a sorting for size to work both ways; if the larger molecule is held back, the smaller can be made to pass, but an aperture cannot hold back the smaller and admit the larger of a pair. To fulfill this need, we use a shunting system. A set of cavities can admit the small molecule with the larger. If this channel is tortuous, so that many collisions occur between molecules and cavity wall, the wall can be pierced with apertures which will pass the smaller but not the larger. Thus if we look only at molecules which have entered the cavities of the membrane, we can adjust the dimensions so that by the time the final aperture which terminates the cavity is reached, the fraction of the smaller molecules is reduced by any desired amount. They have been shunted down side tracks to the appropriate surface of the membrane. But there is of course a difficulty. The side tracks may be terminated by apertures too small to admit the back flow of the larger molecules, but they must still admit the entry of the smaller molecules, if they are to deliver small molecules back to the side whence they start. Thus a back diffusion can begin, and at full equilibrium, the concentration of small molecules at the final aperture will be just the same as that in the bulk solvent, and the membrane fails to work. But now it is clear that we must work infinitesimally away from equilibrium. The rate of diffusion is affected by the speed of motion of the membrane, which moves toward the region of high concentration of the small molecules. If the membrane cavities are opened to the small molecules by apertures which face the direction of motion, and the back-diffusing molecules into the side-tracking channel must enter apertures facing away from the motion, a difference in diffusion rates can be produced. Any finite difference will be enough to give an arbitrary concentration difference at least if we use many layers of membranes with relative velocities, for the geometry is at our disposal. Thus the set of *moving* membranes will work to admit large molecules only, to any degree of separation desired, given a fixed but finite departure from true equilibrium. It is uncertain what this costs in free energy. Perhaps the case is no different from other theoretical processes in thermodynamics if these are carefully examined, though our example is surprisingly cumbersome.

If we accept the above arguments as demonstrating at least the plausibility of discrimination of an arbitrary molecule A from B, with membranes passing either molecular species at our choice, then a discrimination among any finite set of molecules is to be secured by placing the appropriate *go no-go* membranes in series, to form a complex piston which passes only one member out of any arbitrary finite set of distinguishable species. The larger the possible set, the more complex the membrane. We shall later return to employ such devices.

2. THE SCALE OF THERMALLY SIGNIFICANT COMPLEXITY

The subtlety of the visual patterning of living forms is marked in every butterfly and blossom. The behavioral complexity for which the same pair of organisms may stand as paradigm is no less wonderful. Yet the hopefulness of the general program upon which we are embarked is won only at the price of the denial of importance to phenomena of that very class. Nearly all the manifest visual and mechanical intricacy of organisms, like their apt behavior, turns out to be without quantitative thermodynamic importance. Morphology and ecology are from the present point of view only small secondary properties of a fundamentally thermodynamic system, just as the secondary qualities of matter are derived from the interaction of two atomic systems, the sample, and the end-organ of the sensor. It would be an error to hold that this thermodynamically unified characterization in any way reduces the importance or the content of the more detailed biological account of the real living system, just as the reduction of sensed signals to atomic properties is the basis of, but by no means the dismissal of, a successful analysis of perception.

The force of the paragraph above rests on an entirely numerical result.[5] We can make this clear in a simple illustrative example, which will strongly suggest its own generalizability. Consider, for example, a two-dimensional array of n^2 black and white cubes, of edge length L arranged in equal number in a distinctive pattern, one layer thick, over a square nL units on a side. It is not hard to compute the information content of such an array. The usual formulas agree, of course, with a calculation based on the familiar Boltzmann definition of the thermodynamic probability of the array. It is more realistic to allow noise in the pattern; for example, if the pattern were a checkerboard, one could think of the black–white boundaries as diffuse and not as precisely sharp. The contribution of such an array to a free energy is just $-kTS$, and we may write, to a satisfactory approximation,

$$F_{\text{pattern}} = kTn^2 \ln f \, ,$$

where f measures the information contained in each discrimination of a black cube from a white one. The parameter f would be 2 if the array were unique; if it is diffuse, and in the ensemble the pattern varies within the limits of the noisy definition from case to case, the ensemble average of f is somewhat less than 2, though not by a large fraction if the array retains reasonable definition. But the cubes must be made of some material which possesses its own chemical free energy in the conditions of its environment. For portions of organisms we can expect this to be something like kilocalories or tens of kilocalories per mole with respect to a degradation into energy-poor products like water and carbon dioxide. Write the material free energy per gram then as $F_{\text{mat}} = N_0 kTh/\mu$, where h is a number of order one to ten; N_0 is Avogadro's number, μ the gram-molecular weight. Now the ratio of the free energy represented by the choice of pattern to that of the material in which the pattern is necessarily embodied is just

$$F_{\text{pattern}}/F_{\text{material}} = \mu \ln f/\rho L^3 N_0 h \, ,$$

where L^3 is the volume of the cubical unit, whose density is ρ. A small table of the ratio will make the point eloquently, if approximately:

Free energy ratios for patterns built of subunits of edge L ($\mu \cong 100$ g/mole):

	L	F_p/F_m
Visible pattern	$\frac{1}{2}$ mm	$\sim 10^{-18}$
Microscopic pattern	1 micron	$\sim 10^{-10}$
Electron micrographic pattern	0.1 micron	$\sim 10^{-7}$
Molecular pattern	10 Å	$\sim 10^{-1}$

---

[5] This section is merely a gloss on the cogent note of J. A. V. Butler, Nature **158**, 153 (1949). A similar treatment has been given by P. Porcino, M.S. Thesis, Cornell University, 1951 (unpublished).

The tendency of the table is clear: any contribution of the patterning to the thermodynamic functions is likely to be negligible until the pattern is constructed with molecular fineness. It is this of course which ensures that the patterns of artifice, from printing to machinery, can be ignored in the thermodynamics of their realizations.

It is to be observed that we have discussed this question without particular care to show the equivalence of the information-theoretical formulation with that of classical thermodynamics. That topic will be more carefully set forth in the following section. Here we rely wholly upon the orders of magnitude to free us from the need to discuss pattern in terms other than molecular. Visual coupling with organic order, the imputed disassembly of intricate structures above cellular level, and similar points, which can make the relation between pattern and ordinary chemical thermodynamics a knotty one, are avoided by the observation that most such features of a system, even if we assign them their full content of information, cannot appreciably affect the thermodynamic behavior. They retain their importance for more detailed ways of knowing. But in what follows we are freed to discuss systems rather more chemically than would appear to be appropriate for butterflies, mitotic patterns, organelles, or neurons.

3. THE THERMODYNAMICS OF SELF-REPRODUCTION

We shall consider a large volume of material, called the substrate, which undergoes a chemical change. Its initial free energy, reckoned for a fixed volume and temperature, is just the familiar Helmholtz free energy, $F = E - TS$. Only isothermal changes are of much biological importance, and any gases taking part are normally held in solution, so that volume changes and hence any pV work are negligible. Within the volume is the system which undergoes reproduction. Its free energy, like that of the substrate, is measured from a fixed state of the substrate material, called the depletion state. (It is of course invariably necessary in thermodynamics to define the lowest admissible state of the system by some conventional means. This corresponds in practice to fixing the maximum time of observation. No one waits for the spontaneous fusion, for example, of the two protons of water.) The free energy change during the process observed, which we will call self-reproduction, is $-\Delta F_{\text{sub}} = F_{\text{substrate}} - F_{\text{depletion}} + F_{\text{system}}$. The total $-\Delta F_{\text{sub}}$ is necessarily positive, or at best zero. No process can cause the free energy to increase. (See Fig. 1.)

Quite generally then the free energy changes only within the constraints imposed by the available reaction paths. In this present case we are paying attention to the constraint imposed by the need to reproduce that portion of the whole which we have called the reproducing system. We shall say that the reproductive process is required to produce a copy of the system. But what is meant by a copy? We do not in general expect atom for atom replication. Rather we impose a test in principle fully empirical. Among the many alternatives which we might find for the system we will regard reproduction as having taken place if the new system—call it the copy—passes the test. The test can be applied, for example, by requiring that the copy in fact pass a set of semipermeable membranes which are capable of filtering out the system itself from every other member of the class of possible alternatives. Such tests differ widely, depending on the process under study. The number of possible alternatives may be small, and the test therefore adequate even if simple; or the number very large, and the test therefore extremely complex. But whatever test is assigned, the membranes used must have been selected from some set, and in the simplest case, with equal weights from an ensemble of N alternatives. Then the free energy assignable to the test is just $F_{\text{test}} = kT \ln N$. We divide this by the free energy change ΔF_{sub}, to form the dimensionless parameter

$$\beta(\text{for } \beta\iota o\varsigma) = F_{\text{test}}/\Delta F_{\text{sub}} \, .$$

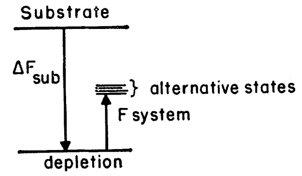

Fig. 1. The free energy changes involved in the parameter β.

Consider the values which β may assume. Evidently it ranges over the domain $0 \leqslant \beta \leqslant 1$, for the free energy ΔF_{sub} must be positive, and since at most the mixing entropy separation made by the membrane set could have stored *all* the available free energy, $\beta \leqslant 1$. For the other limit, consider a saturated solution as substrate out of which a crystal forms. If we know the chemistry well enough to be sure that only a single type of crystal is possible, there is no need for any "test." A mere mass detection would allow the "self-reproduction"—here dwindled to a simple spontaneous process—to be recognized. Such a case corresponds to a value of β in the neighborhood of zero. A more instructive case is the formation of an optically active crystal by the growth of a solid phase on a seed of given rotatory power, placed in a saturated solution of racemic type. The free energy of solution is some RT per gram, say. Now the test of self-reproduction is a measure of optical activity, which has to be done with some accuracy if we are to insist that the new crystal of a gram or so is really all—or nearly all—of the proper enantiomorph. Now it is clear that the test permits alternatives in the number given by its discrimination: if it measures to one percent accuracy we expect to select a replica out of some hundred. Then $\beta \cong (kT \ln 100)/(\text{few } RT) \cong 10^{-22}$. The familiar metaphor which relates crystal growth to the living process is here given some quantitative force: a simple spontaneous growing of a crystal has very small stature in the scale; a more discriminate one, in which the "seed" plays a genuine role beyond mere kinetics, acquires a larger value of β. It is easy to imagine, though not so easy to exemplify, an auto-catalytic process, in which a "seed" catalyzes a very specific polymer to form the crystal, out of a mixture of many monomers. Such a system would be closer to what we mean by life; its β would plainly be larger; it would fully satisfy the older qualitative chemical definitions of living system. By this stage, we are ready to consider such systems as the Kornberg preparation, in which primer DNA is replicated. The inadequacy of the genetic characterization of the replicated material, and the relatively low yield of the process, suggest that β does not rise very high. Rough estimates would place β somewhere below $10^{-(3 \text{ or } 4)}$. If the yield approached complete use of the monomer nucleotides, and the material were genetically functional, β could be a few percent, and it would be hard to deny the system full kinship with virus growth. Virus growth itself, even cellular replication, does not seem to come very close to $\beta = 1$, if our crude guesses as to the size and specificity of the genetic material on the one hand, and the free energy consumption on the other are anywhere near correct.[6] But somewhere along the scale of increasing β, between a value below the Kornberg preparation, say at 10^{-4}, and one above the virus, say with values near 10^{-2} or 10^{-1}, the system became alive by consensus.

Man-made systems, ruled by the necessities of scale, come nowhere close as yet. The ingenious mechanical models, like that employing floats bearing electrets and electromagnets described by Morowitz,[7] are *literally* self-reproducing. A structure AB will select from an environment of free-floating A units and B units one of each, catalyze their combination into a new AB and set it free under Brownian collisions to multiply and replenish the species. But the value of β is lamentably small, below 10^{-18}! The feeling that such systems are relevant, but inadequate, models of life is confirmed by the parametric scale with complete satisfaction. The usual less precise models, say printing processes, molding, and the rest, come even more clearly under the same stricture. The final metaphor we shall mention, the kindling of a flame on an appropriate substrate of candle by a candle flame itself, is clearly less suited to the sort of description of chemical systems here made the basis, but it is plausible to say that a similar approach would assign to this process also a very low value of the key parameter β.

It should be observed that the fact that the various copy states might themselves have different material free energy content has been ignored. Any effort to weigh this free energy—the full F_{system}—into the balance with the *essentially nonthermal free energy* of F_{test} is likely to produce paradox. For then any system which in reproduction merely drops a small part of the way toward the depletion state will appear to have a high value of β. For example, a system yielding ethyl alcohol from sugar in water would retain much free energy as long as the alcohol remained intact. If the test were alcohol recognition, but the free energy available upon removing the test constraint counted as the numerator of β, a very plausible definition, then β would be high. But if, as we have taken it, only discrimination among the alternatives which include ethyl alcohol—a rather small set of molecules below a molecular weight say of a few hundred—contributes to β, then this system finds a very small value as well.

The parameter β seems to characterize a wide variety of systems, including most of those which have provided the counterexamples to most of the definitions of life, as well as most of those which have always been suggestive metaphors for life. But the values of beta clearly part the examples into two classes, low values, which are plainly nonliving, and values in the rough neighborhood of one, which are reached as far as we can see only by processes embodying that marvelous intricacy of the living stuff.

4. THE SPONTANEOUS ORIGIN OF LIVING PROCESSES

This continuum representation of the transition from trivial to living self-reproduction has bearing on the description of the actual transition which took place at least once in the past. Consider any

---

[6] Compare C. F. Ehret, in *Information Theory in Biology*, edited by H. P. Yockey (Pergamon Press, Inc., New York, 1958), p. 218, and other papers in that volume.

[7] H. J. Morowitz, Am. Scientist **47**, 261 (1959).

natural environment; for simplicity (not to be taken literally), we will call it sea water. In a sea-water volume, even if there were no current of incoming free energy from the sun, thermal fluctuations will produce a population of molecular forms which differ in free energy content from the minimum available. This is described usually by the familiar form[8]

$$N \propto \exp\left(-\Delta F/kT\right) .$$

Of course the presence of a free energy input will only strengthen the argument, though it may modify the functional form of N. In the early sterile environment, as Darwin first[9] said in that prescient letter of 1876, molecules will arise, to be sure in small numbers, with any degree of complexity, as measured by the free energy deviation ascribable to their form, $\beta \Delta F_{\mathrm{sys}}$. Each such molecule is present only transiently, though the population is in a steady state. If external free energy is supplied, the effect is roughly to be represented by assigning an effective temperature, say T^*, instead of the kinetic temperature T. Since no fast-acting specific catalysts are present to reduce the lifetime of the complex states, they accumulate as described. But it can now happen that a few such states, at the far end of the scale of complexity, rare but interesting deviates from expected free energy minima, have some auto-catalytic effect. For example, a molecule A occurs which produces another B, rather often from the environment, at the price of a free energy reduction, and induces a depression in the population just below the complexity level of B. Then the pair A and B might combine to reproduce themselves, again spending free energy, and making now a real depletion of their near-lying precursors. This process evolves, and sooner or later a real population gap occurs between the self-reproducing complex end and the free-energy depleted end of the molecular distribution. Life has begun.

It seems plausible that this process provides a more general account of the evolution of life from the nonliving by easy stages than is available from biochemistry of the present molecules, ATP, DNA, RNA, proteins, and the like. Other chemical beginnings do not appear thermodynamically at all different. One might expect a wide variety of fundamental self-reproducing systems. Moreover, the early beta values were doubtless less spectacular ones than those now possessed by life. This could occur for two reasons: (i) the earliest intermediates were nearly the same as the self-reproducing molecules so that the biochemical reaction chains evolved backwards to use of the much simpler intermediates, very frequent in the population, just as Horowitz[10] long ago pointed out, (ii) the self-reproduction was sloppy, by no means producing progeny so well-characterized as do modern cells or viruses. Life itself thus evolves along the scale of β to form the clear gap in population now present. In Fig. 2 this succession of events is schematically presented. A major question which it may be given to the next decade to solve is whether there is life on Mars, and whether that life whose reality is suggested, though not proved, by our present knowledge, is related to terrestrial life only in general and thermodynamical ways—large beta, mutation, evolution, common energy sources, helical polymers, optical activity—or also by biochemical similarities.

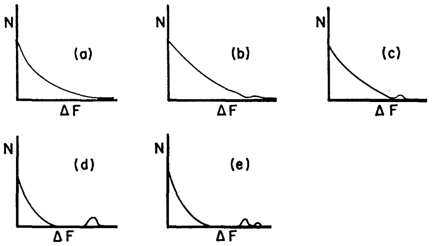

Fig. 2. The time sequence of changes in a molecular population arising from fluctuations leading to the presence of self-reproducing systems.

Finally, it is tempting to see in the hazy future of ultra-microelectronics,[11] as well as in the clearer program of molecular biology, the rise of a second peak of large beta values even beyond those of ordinary life. This would be man-made, or "second" (maybe third) life, with an information-theoretic but possibly neither a chemical nor a physical continuity with ordinary "first life." For such systems, an approach beginning with a point of view like that here expressed will be worth completing. The heuristic effort at definition is made finally valid by innovation; there may yet be new things under the sun, which we shall need to describe.

---

[8] R. C. Tolman, *Principles of Statistical Mechanics* (Clarendon Press, Oxford, 1938), p. 638.

[9] Cited in part in F. Darwin, *Life and Letters of Charles Darwin* (Appleton, New York, 1886), Vol. II, p. 202.

[10] N. H. Horowitz, Proc. Natl. Acad. Sci. U. S. **31**, 153 (1945).

[11] Consult the reports of the remarkable work-in-progress by K. R. Shoulders at Stanford Research Institute, California (unpublished.)

5. ACKNOWLEDGMENTS

All my years as a physicist have been informed by what Professor J. Robert Oppenheimer, as teacher, friend, and Director, has given me. I hope the spirit of those past years will be evoked if I mention in gratitude also our old Pasadena friend, Richard Tolman. It is a pleasure to recall more recent discussions with my associates Ethel and Irwin Tessman of Purdue University, Malcolm Skolnick of the Institute for Advanced Studies, David Hawkins of Elementary Science Study and the University of Colorado, and V. F. Weisskopf of CERN. The quality of their help, like that of my old teachers, is fairly to be appraised not by what I have written, but by these earnest thanks.

Structure, Substructure, and Superstructure*

CYRIL STANLEY SMITH

Massachusetts Institute of Technology, Cambridge, Massachusetts

Anyone who works with the microscope for an intellectual or a practical purpose will frequently pause for a moment of sheer enjoyment of the patterns that he sees, for they have much in common with formal art. What follows is an attempt to extend into a more general field some views on the nature of organization and relationships that arose during many years of study of the microstructures of metals and alloys.[1] In a landscape painting of the Far East, a rock in the foreground with cracks and crystalline texture is often echoed in a distant mountain with cliffs, chasms, wrinkles, and valleys; a tree may be related to a distant forest or a turbulent and eddied stream to a distant tranquil pond. Each part with its own structure merges into a structure on a larger scale. Underlying structures are imagined as a necessary basis for the visible features. The connectivity of all is suggested by the branching treelike element of the design. Both separateness and continuity are interwoven, each necessary to the other and demonstrating the relationship between different features on a single scale and between units and aggregates on differing scales. There is an analogy between a work of art which suggests an interplay of dimensions and the real internal structure of a piece of metal or rock which results from physical interactions between the atoms and electrons composing it.

The study of microstructure on the scale within the range of the optical microscope (dimensions between a micron and a millimeter) is a somewhat old-fashioned branch of science, and it still involves a high degree of empirical observation and deduction. Far more "highbrow" is the rigorous science and simple elegant mathematics of the ideal crystal lattice considered as point groups in space. The whole field of crystal structure, mathematically developed in the nineteenth century by Bravais, Federov, and Schoenflies, was experimentally opened up by Von Laue and especially the Braggs in 1912–13, using the diffraction of x rays to reveal and to measure the periodicities and symmetries in the arrangement of planes of atoms in crystals. But the mathematical physicist must simplify in order to get a manageable model, and although his concepts are of great beauty, they are austere in the extreme, and the more complicated crystal patterns observed by the metallurgist or geologist, being based on partly imperfect reality, often have a richer aesthetic content. Those who are concerned with structure on a superatomic scale find that there is more significance and interest in the imperfections in crystals than in the monotoous perfection of the crystal lattice itself. Like the biologist, the metallurgist is concerned with aggregates and assemblies in which repeated or extended irregularities in the arrangement of atoms become the basis of major structural features on a larger scale, eventually bridging the gap between the atom and things perceptible to human senses.

---

* Some readers of the *Reviews of Modern Physics* may feel affronted by the appearance of this elementary, seemingly naive, paper in their journal. But others may see a certain appropriateness, and may even enjoy this attempt to share with nonscientists the visual pleasures of the laboratory. Certainly, it would not have been written if its author had not for some years been exposed to Robert Oppenheimer's sensitive view of the world. The paper, with more illustrations, will appear in the *Vision and Value* series, a collection of essays on the contemporary scientific and artistic environment, edited by Gyorgy Kepes. This will be published by George Braziller, Inc., who have granted permission for the present publication.

[1] The converse relationship between aesthetics and metallurgy—the influence of the techniques discovered by craftsmen making works of art upon the development of the science of metals—was discussed at some length in C. S. Smith, *A History of Metallography* (University of Chicago Press, Chicago, 1960).