Available online at www.sciencedirect.com
SCIENCE DIRECT®
Journal of Theoretical Biology 238 (2006) 949–961
Journal of Theoretical Biology
www.elsevier.com/locate/yjtbi

Organizational invariance and metabolic closure: Analysis in terms of (M, R) systems

Juan-Carlos Letelier[a,*], Jorge Soto-Andrade[b], Flavio Guíñez Abarzúa[b], Athel Cornish-Bowden[c], María Luz Cárdenas[c]

[a]*Departamento de Biología, Facultad de Ciencias, Universidad de Chile, Casilla 653, Santiago, Chile*
[b]*Departamento de Matemáticas, Facultad de Ciencias, Universidad de Chile, Casilla 653, Santiago, Chile*
[c]*Bioénergétique et Ingénierie des Protéines, Institut de Biologie Structurale et Microbiologie, Centre National de la Recherche Scientifique, 31 chemin Joseph-Aiguier, 13402 Marseille Cedex 20, France*

Received 21 February 2005; received in revised form 5 July 2005; accepted 7 July 2005
Available online 24 August 2005

Abstract

This article analyses the work of Robert Rosen on an interpretation of metabolic networks that he called (M, R) systems. His main contribution was an attempt to prove that metabolic closure (or metabolic circularity) could be explained in purely formal terms, but his work remains very obscure and we try to clarify his line of thought. In particular, we clarify the algebraic formulation of (M, R) systems in terms of mappings and sets of mappings, which is grounded in the metaphor of metabolism as a mathematical mapping. We define *Rosen's central result* as the mathematical expression in which metabolism appears as a mapping f that is the solution to a fixed-point functional equation. Crucially, our analysis reveals the nature of the mapping, and shows that to have a solution the set of admissible functions representing a metabolism must be drastically smaller than Rosen's own analysis suggested that it needed to be. For the first time, we provide a mathematical example of an (M, R) system with organizational invariance, and we analyse a minimal (three-step) autocatalytic set in the context of (M, R) systems. In addition, by extending Rosen's construction, we show how one might generate self-referential objects f with the remarkable property $f(f) = f$, where f acts in turn as function, argument and result. We conclude that Rosen's insight, although not yet in an easily workable form, represents a valuable tool for understanding metabolic networks.

Keywords: (M, R) systems; Metabolic network; Metabolic closure; Infinite regress; Systems biology

1. Background

The massive, extended and often cryptic scientific output of Robert Rosen poses a scientific dilemma. While the great majority of biologists are unaware of his work, a few regard it as being of the kind to be expected once in a thousand years,[1] and a few others are trying to bring understanding of it to the point where its importance for biology can be objectively assessed.[2] An essential step in gauging its relevance is to understand the core of Rosen's thinking, which concerns a formal theory of metabolic networks and the notion of circularity or metabolic closure. For this reason we shall concentrate here on his investigations of metabolic systems and his definition of living systems; we shall not

*Corresponding author. Fax: +56 2 271 2983.
E-mail address: letelier@uchile.cl (J.-C. Letelier).

[1]A measure of this admiration can be seen in a recent paper with the unexpected title of "Robert Rosen (1934–1998): a snapshot of biology's Newton" (Mikulecky, 2001). More realistically, the mathematician John Casti recently said "The work of Rosen will keep scholars busy for decades" (Casti, 2002).

[2]Rosen summarized his work in his opaque but important book *Life Itself: a Comprehensive Inquiry into the Nature, Origin and Fabrication of Life* (Rosen, 1991).

0022-5193/$ - see front matter © 2005 Elsevier Ltd. All rights reserved.
doi:10.1016/j.jtbi.2005.07.007

touch on some other aspects of his work, such as his epistemological research and his definition of *complex systems*.

The central motto of Rosen's research programme ("Organisms are closed to efficient causes") (Rosen, 1991) can be traced to an insight on the nature of cellular metabolic networks (Rosen, 1958a,b, 1959); this consists of a semi-formal method to explain how the network of biochemical processes that constitutes metabolism *bootstraps* itself without the help of external agents generated outside the network, thus keeping cell organization invariant in spite of continuous structural change. He based a large part of the development of his ideas on a branch of mathematics known as Category Theory, and in mathematical terms his major insight appears as a peculiar result about sets and admissible transformations between them. Rosen appears to have been aware of the rather peculiar nature of this result and he gave several slightly different proofs with slightly different interpretations (Rosen, 1958b, 1959, 1966, 1967, 1971, 1972, 1991, 2000); maddeningly, however, he never provided any concrete examples, not even mathematical ones, let alone ones that would be intelligible to biologists. This result, here called *Rosen's Central Result*, is, without doubt, the core of his view of theoretical biology, as well as the unavoidable starting point for analysing his views on complexity.

The central result has been used recently in attempts to expand his ideas to other areas, such as bioinformatics (Wolkenhauer, 2001, 2002), control theory (Casti, 2002), and even sociology (Nomura, 2002). Although these recent publications seem to imply increasing acceptance of Rosen's ideas, we must point out that recently his entire approach has been called into question and even declared "false" (Landauer and Bellman, 2002). Because of this, and because it is such a special result, with important implications for theoretical biology and computer science, we have found it necessary to revisit and clarify it, trying to give some examples and connect it to other theoretical ideas. In doing so, we have changed somewhat the nomenclature and notations that he used. Renaming established terms is not, of course, something to be done lightly, but we believe it to be unavoidable in this instance because Rosen used a number of words, such as *replication*, in ways that can only confuse readers familiar with their usual meanings in biology. We shall, however, take care to be explicit whenever we alter one of his terms, and will note why we regard it as unsatisfactory.

The structure of this paper is as follows: in Sections 2 and 3 we introduce (M,R) systems, in Section 4 we enunciate the central result, and in Section 5 we set out its mathematical context. In Section 6 we show two examples of (M,R) systems, and in Section 7 we introduce what we consider a possible generalization of Rosen's result about infinite regress and closure. Finally, in Section 8, we discuss the many implications of Rosen's ideas for the study of metabolic networks. Although category theory was central to Rosen's thinking, we do not use it here because it is this that gives much of his writing its abstract character, making it opaque to most biologists. Nonetheless, a full understanding of his work requires some appreciation of what categories are and how they are relevant to his analysis of metabolism (Pierce, 1991; Joslyn, 1993). Some of these ideas have been briefly addressed in a previous article (Letelier et al., 2004).

2. Introduction to (M,R) systems

Rosen's basic ideas appeared initially in three papers (Rosen, 1958a,b, 1959), where he introduced a formal model of metabolic networks that he called (M,R) systems. These systems owe many of their properties to graph theory and the 1950s-style black-box analysis of electronic circuits, but (more crucially) to an interpretation that identified an enzyme with a mathematical mapping. Grasping the details of the formal structure of (M,R) systems is an essential first step in understanding Rosen's ideas on circularity. In our description we have found it necessary to change the original nomenclature as detailed in Table 1.

2.1. The M components

The theoretical model that Rosen applied to metabolism and metabolic networks starts with a simple formalization of biochemical reactions. According to him every metabolic reaction, such as the one catalysed by the enzyme glucokinase:

$$\text{Glucose} + \text{ATP} \longrightarrow \text{Glucose 6-phosphate} + \text{ADP}$$

can be formalized as

$$a_1 + a_2 \longrightarrow b_1 + b_2. \tag{1}$$

This process can be viewed as the action of an operator M that transforms molecules a_1 and a_2 into b_1 and b_2:

$$a_1 + a_2 \overset{M}{\longrightarrow} b_1 + b_2, \tag{2}$$

Table 1
Comparison of Rosen's terminology with that used in this paper

Rosen's terminology	Terminology used in this paper
Component	Catalyst (or enzyme)
Repair	Replacement
Replication	Organizational invariance
Replicative (M,R) system	Organizationally invariant (M,R) system
Transformable molecule	Metabolite

here M behaves like an enzyme, or, more generally, a catalyst, but Rosen called these operators *components*, whereas for a_1, a_2, b_1 and b_2, which correspond to metabolites in normal biochemical usage, he used the term *transformable materials*. Thus, a component transforms input materials into output materials according to:

$$\textit{Input materials} \xrightarrow{M} \textit{Output materials}. \tag{3}$$

The catalyst M, therefore, acts formally as a mathematical mapping, because it transforms some variables (from the admissible set of input materials) into some variables belonging to the set of admissible output materials. As enzymes are not totally specific for the types of molecules that they transform, Rosen interpreted M as a mapping between two sets defined by Cartesian products:

$$M : A_1 \times A_2 \longrightarrow B_1 \times B_2,$$
$$(a_1, a_2) \mapsto M((a_1, a_2)) = (b_1, b_2). \tag{4}$$

From this point of view the catalyst M can accept for a given input not only $a_1 \in A_1$ but also molecules $a_1', a_1'', a_1''', \ldots, \in A_1$ that are similar to but not identical to a_1. With this model the action of an enzyme can thus be framed in the language of mappings, as M acts as a particular mapping between sets $(A_1 \times A_2)$ and $(B_1 \times B_2)$.

Although Rosen did not mention it, this over-reaching formalization is extreme. An enzyme can be presented in vitro with artificially produced molecules that are accepted and processed as substrates because of their structural resemblance to the natural substrate, and it then appears that the set A_1 is "large". However, matters are radically different in vivo because in the organism only one (or a few) acceptable substrates exist. For example, in some bacteria the enzyme glucokinase mentioned above will not accept any natural sugar substrate other than glucose. Higher organisms typically use somewhat less specific enzymes known as hexokinases to catalyse the same reaction, but even these accept only a small range of sugar substrates, typically including mannose and fructose but not galactose (Cárdenas et al., 1998).

Using this mathematical model for a single metabolic reaction Rosen generalized it to take account of the complete network of biochemical reactions that constitute a living metabolism. Consider the following (simplified) metabolic network composed of three elementary reactions, which correspond to the very simple metabolic network shown in Fig. 1:

$$M_1 : A_1 \times A_2 \longrightarrow B_1 \times B_2,$$
$$(a_1, a_2) \mapsto M_1((a_1, a_2)) = (b_1, b_2), \tag{5}$$

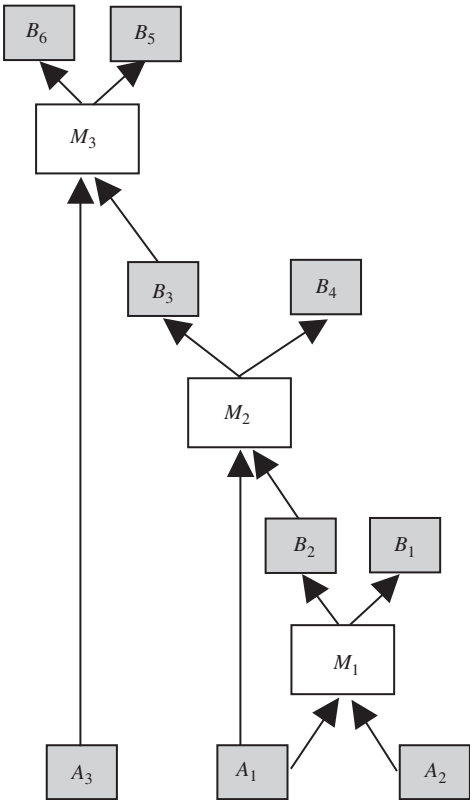

Fig. 1. An initial view of a metabolic network as a collection of catalysts (*components* in Rosen's terminology: see Table 1) (M_1, M_2, M_3) that transform metabolites from specific sets $(A_1, A_2, A_3, \ldots)$ into specific sets $(B_1, B_2, B_3, \ldots)$.

$$M_2 : A_1 \times B_2 \longrightarrow B_3 \times B_4,$$
$$(a_1, b_2) \mapsto M_2((a_1, b_2)) = (b_3, b_4), \tag{6}$$

$$M_3 : A_3 \times B_3 \longrightarrow B_5 \times B_6,$$
$$(a_3, b_3) \mapsto M_3((a_3, b_3)) = (b_5, b_6). \tag{7}$$

Rosen interpreted the simultaneous action of M_1, M_2 and M_3 as the action of a generalized catalyst M_{met} defined between the following sets (Fig. 2):[3]

$$M_{met} : A_1 \times A_2 \times A_3 \longrightarrow B_1 \times B_2 \times B_3 \times B_4 \times B_5 \times B_6,$$
$$(a_1, a_2, a_3) \mapsto M_{met}((a_1, a_2, a_3)) = (b_1, b_2, b_3, b_4, b_5, b_6). \tag{8}$$

If we apply the same procedure for the complete metabolism of a real organism, composed of thousands of steps, we can interpret the overall metabolism M_{met} as

[3]Other approaches are possible, such as:

$$M_{met} : (A_1 \times A_2) \times (A_1 \times B_2) \times (A_3 \times B_3)$$
$$\longrightarrow (B_1 \times B_2) \times (B_3 \times B_4) \times (B_5 \times B_6)$$

This is more than a notational difference, as the application of Rosen's ideas to real metabolic networks would certainly demand changes in his initial formulation.

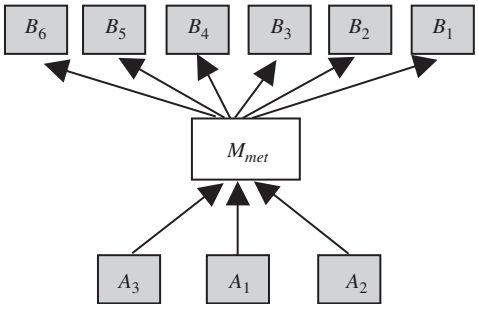

Fig. 2. The network of Fig. 1 viewed as the action of a single catalyst (M_{met}). All the intermediate steps have disappeared, though the intermediate metabolites (B_2 and B_3) have not.

the following mapping:

$$M_{met} : A = (A_1 \times A_2 \times A_3 \times \cdots \times A_p) \longrightarrow B = (B_1 \times B_2 \times \cdots \times B_q),$$
$$\mathbf{a} = (a_1, a_2, a_3, \ldots, a_p) \mapsto M_{met}(\mathbf{a}) = (b_1, b_2, \ldots, b_q) = \mathbf{b}. \quad (9)$$

In a very compact notation, therefore, we can write the complete metabolism as the (huge) mapping M_{met} between the (huge) sets A and B.[4]

$$M_{met} : A \longrightarrow B, \quad \mathbf{a} \longmapsto M_{met}(\mathbf{a}) = \mathbf{b}, \quad (10)$$

where M_{met} represents the action of the overall metabolism and can be interpreted as a type of *generalized enzyme* that transforms one instance of input molecules $(a_1, a_2, a_3, \ldots, a_p) \in A$ into an instance of output molecules $(b_1, b_2, \ldots, b_q) \in B$. As many metabolisms are theoretically possible between sets A and B we define $\mathscr{M}$ conceptually as the set of all possible metabolisms between A and B.

Does the set $\mathscr{M}$ have a structure? After all, a metabolic network is much more than just a random collection of transformations between molecules. As we will see, this is a crucial point that was never clarified by Rosen. In his original arguments he assumed that

$$\mathscr{M} = Map(A, B) = \text{set of all mappings between sets } A \text{ and } B.$$

We shall see that this identification is too general, and that $\mathscr{M}$ must be a proper subset of $Map(A, B)$. Up to this point Rosen's contribution has been to use the language of *mappings* to interpret the overall metabolism in the context of *graph theory*, with the added complication of considering the sets of intermediate metabolites. If this had been his main contribution, it would have only anticipated, without exploiting, some of the approaches developed in metabolic control analysis (Hofmeyr and Cornish-Bowden, 2000).

2.2. The R subsystems

Rosen's crucial insight concerns the long-term stability of a metabolic network. Because catalysts $M_1, M_2, \ldots$ etc. represent physical objects (enzymes, in the case of cellular metabolism) they must be subject to wear and tear, i.e. they must inevitably become degraded by a wide variety of processes. Every M_i has only a finite lifespan before disappearing, and if the overall metabolic network representing an autonomous cell must continue operating in a steady state it must be replaced as fast as it is degraded. A large part of enzyme degradation in practice is specifically catalysed by enzymes and is not simply due to protein lability. However, although this is important it does not affect the main points to be discussed; indeed, rather than simplifying the problem of infinite regress that we shall consider, it greatly complicates it. But what elements (or reactions) allow such replacement? Rosen posited that for every catalyst M_i there must exist a subsystem R_i, made of subnetworks of biochemical reactions, that uses intermediate molecules, from the set B, to replace M_i or, in Rosen's terminology, to "repair" it (Fig. 3).[5]

This insight is essential for understanding the biological relevance of Rosen's work. In contrast to a man-made machine, for which worn-out components can be replaced by an agency external to the machine itself, a living organism is a self-made machine, and all of its catalysts must be made and, when necessary, replaced within the system by catalysts that are themselves products of the system, are themselves degraded, and must also be replaced, again by components within the system, that are likewise degraded, and need to be replaced, and so on indefinitely. In such a system the possibility of infinite regress is obvious, and Rosen's work can be interpreted as a search for a way to escape this, or in other words a search for *closure*.

Although a highly sophisticated machine may contain gauges to warn its operator that certain parts are worn out or faulty and need to be replaced, no machine provides such information about *all* of its components, and even if it did it would still need to rely on an external agency to carry out the actual replacement. But a completely autonomous living organism needs to encode all of the information about the state of all of its catalysts, and,

[4]Here we are departing slightly from Rosen's texts. In his papers he viewed a as an input and b as an output. We view a as the collection of left-hand sides of biochemical equations, and b as the collection of right-hand sides.

[5]To *maintain by replacement* is the real concept behind the notion of repair. The word *repair* was not a very happy choice, and in this paper we use the term *replacement*.

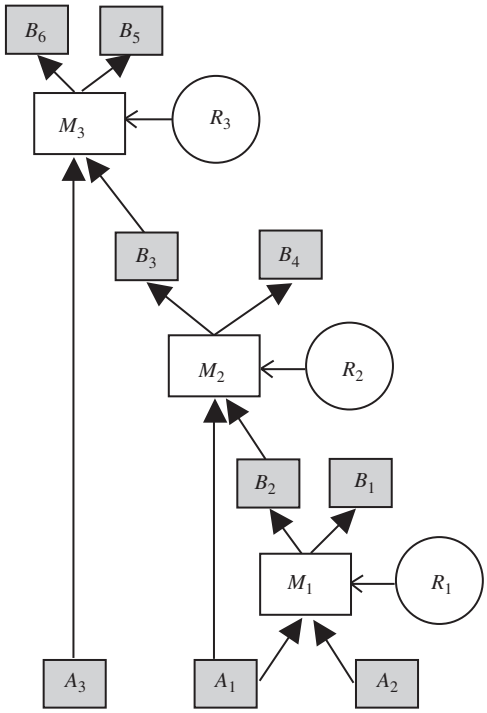

Fig. 3. A small (M, R) system with three enzymes ("components") (M_1, M_2, M_3) and three "repairers" (R_1, R_2, R_3). Associated with each M_i is a subsystem (a collection of processes) R_i that maintains by replacement a functional amount of each M_i. Each subsystem is specific to a particular enzyme. Rosen called this process *repair*, but the term *replacement* indicates more clearly what is involved. For simplicity, we treat enzyme degradation as an uncatalysed and unavoidable process, though in real organisms much of it is catalysed.

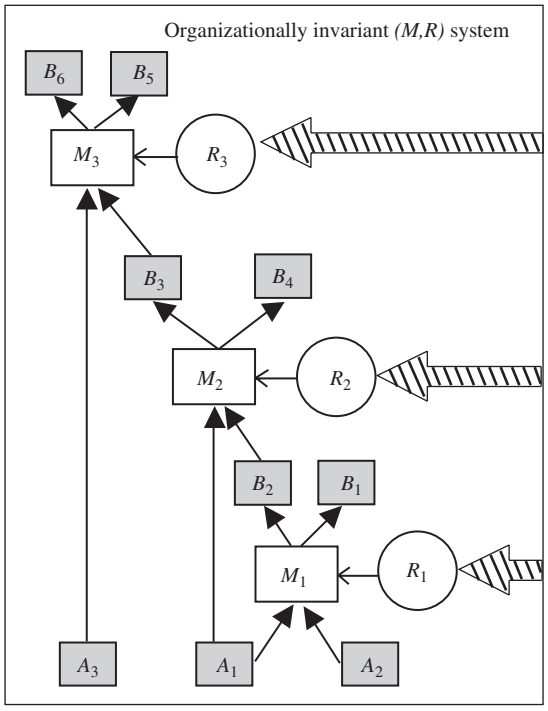

Fig. 4. Logical structure of an organizationally invariant (M, R) system ("replicative (M, R) system"). In this system the replacement of each R_i is a systemic property that depends on the system considered as a whole. The hatched arrows arriving at each R_i, and departing from the system's border, represent this dependence. The proof that this dependence is possible constitutes Rosen's central result.

when necessary, make the necessary replacements itself. According to Rosen, the only way this information can be encoded is in the connectivity of the network itself, i.e. the network must be constructed in such a way that all of its connectivity is uniquely encoded: knowledge of the products of all the reactions should allow the substrates and catalysts to be deduced. Designing a machine to work in this way is vastly more difficult than is recognized in almost any current writing about the nature of life. A recent article of Barbieri (2005), for example, which argues that life is essentially "artifact-making", describes some of the same characteristics of living systems that we mention here, but does not address what we see as the fundamental problem of organizational invariance.

Rosen's insight is an attempt to understand how a system must be organized if it is to continue in operation indefinitely. The crux of the matter is thus to understand how the R_i are replaced while avoiding infinite regress. But, as we have seen, the same wear-and-tear argument that was applied to M_i applies equally well to the R_i. It is possible, of course, but not elegant or useful, to invoke a second-level repair enzyme to replace each R_i. But this "solution" just raises the new question of how to replace the second-level repair enzymes, and is thus no solution at all. Rosen's main result was an insight about the systemic nature (i.e. a property that depends on the connectivity of the network) of this maintenance or "replication" function.[6] Thus, in some (M, R) systems, the total network regenerates all of the R_i: these systems are the (M, R) *systems with organizational invariance* (Rosen called these "replicative (M, R) systems"),[7] and they constitute the model of living systems for Rosen. Remarkably, he found a mathematical expression of this interpretation, and this constitutes his central result, which depends on the mathematical properties of functions and sets[8] (Fig. 4).

3. Algebraic representation of (M, R) systems

In the formalism of Section 2.1, the collective action of the thousands of catalysts M_i is represented by a

[6] *Replication* was another unfortunate choice of term, evoking ideas of biological reproduction whereas the essential notion is *organizational invariance* (i.e. the network maintains its connectivity) under continuous structural change.

[7] (M, R) systems without organizational invariance would be transient entities.

[8] In his initial papers Rosen also obtained some generic systemic laws, such as the necessary existence of at least one catalyst M_c that if destroyed would imply the destruction of the complete network of processes. These results are a direct application of the use of graph theory to (M, R) systems and are not central to the theory.

single mapping M_{met}. Following Rosen, we refer to M_{met} as $f \in \mathscr{M} \subset Map(A,B)$.[9] But how can the collective action of subsystems R_i be represented in the context of mappings? The replacement mechanism is a procedure, denoted by Φ, that, starting with $b \in B$ as input, produces f according to:

$$\Phi(b) = f \quad \text{with the condition}$$
$$b = f(a) \ \text{(for some } a \in A). \tag{11}$$

Because the net effect of Φ is to select from the relatively large set $\mathscr{M} \subset Map(A,B)$ the unique f that fulfills this last equation, using $b \in B$ as an input, it is called a *selector*. Thus, as f represents metabolism, Φ represents replacement. As with metabolism, the selector Φ can also be represented by a mapping between the sets of metabolic configurations (B) and the set of possible metabolisms ($\mathscr{M}$). Again, Rosen assumed the most general structure for the set of selectors $\mathscr{S}$.

$$\Phi \in \mathscr{S} = Map(B,\mathscr{M}) = Map(B, Map(A,B)),$$
$$\mathscr{S} = \text{set of selectors between } B \text{ and } \mathscr{M}. \tag{12}$$

With these definitions it is possible to specify Φ as a mapping with the following properties:

$$\Phi : B \longrightarrow Map(A,B),$$
$$b \longmapsto \Phi(b) = f \ \text{(with } b = f(a) \text{ for some } a \in A),$$
$$\Phi \in Map(B, Map(A,B)) = \text{set of possible selectors.} \tag{13}$$

Then an (M,R) system has the following algebraic structure based on two mappings (f,Φ) acting in synergy:

$$A \xrightarrow{f} B \xrightarrow{\Phi} Map(A,B),$$
$$a \longmapsto f(a) = b \longmapsto \Phi(b) = f.$$

Now, in the full language of maps, we can rephrase the closure result sought by Rosen. How can the selector map Φ be produced by the network when the system is capable of organizational invariance (a replicative (M,R) system in Rosen's terminology), without implying infinite regress?

4. Summary of Rosen's central result

Rosen's solution to avoid infinite regress,[10] was to posit the possible existence of a formal entity β with the property that $\beta(f) = \Phi$. On purely formal grounds, therefore, β is a mapping between $Map(A,B)$ (the set $\mathscr{M}$ of possible metabolisms) and $Map(B, Map(A,B))$ (the set $\mathscr{S}$ of possible selectors). Then β acts as a procedure that, given a metabolism f, produces the corresponding selector Φ that selects metabolism f from $\mathscr{M}$. Thus, without any doubt, β is a rather subtle and difficult entity, because for β to exist it is required that the equation $\Phi(b) = f$, for Φ must have *one and only one solution*, a very demanding constraint. Rosen was unable to produce a clear-cut mathematical description or an algorithm to calculate β; he only showed that its existence was logically possible (see Rosen, 1959, 1972, 1991), and sometimes he recognized that its existence was mathematically difficult (Rosen, 2000, pp. 261–265). The beauty of the concept of β is that it is in some sense "equivalent" to an element $b \in B$, in the sense that β sends f to the *unique* Φ such that $\Phi(b) = f$. It is in this sense that Rosen claims that his construction solves the problem of infinite regress.

The operation of an organizationally invariant (M,R) system can, therefore, be viewed as three mappings (f,Φ,β) acting in synergy:

$$A \xrightarrow{f} B \xrightarrow{\Phi} Map(A,B) \xrightarrow{\beta} Map(B, Map(A,B)),$$
$$f(a) = b,$$
$$\Phi(b) = f,$$
$$\beta(f) = \Phi \quad \text{with } \beta \text{ equivalent to } b. \tag{14}$$

Up to this point we have introduced (M,R) systems (with or without organizational invariance) as Rosen did in his writings (specifically in Rosen, 1958b, 1959, 1972, 1991). In this brief presentation, as in all of Rosen's writings, the precise nature of β has been left open. As β encapsulates the notion of metabolic closure, we clarify the nature of β in the next section.

5. The nature of β and the conditions in which Rosen's central result is valid

Rosen unfortunately never defined the specific conditions under which his central result is valid. Later researchers have also left these conditions unexplained. In this section we clarify some of these conditions as well as the nature of β, which has also remained obscure. To achieve this it is necessary to express Rosen's central result mathematically, and to state and clarify a mathematical proposition involving evaluation maps.

5.1. Generic properties of evaluations

Notation. *For any sets X and Y, we denote by $Map(X,Y)$ the set of all possible mappings from X to Y and by $H(X,Y)$ a given subset of $Map(X,Y)$. For any $x \in X$ we denote by Ev_x the mapping from $H(X,Y)$ to Y which evaluates every mapping $f \in H(X,Y)$ at x:*

$$Ev_x : H(X,Y) \longrightarrow Y,$$
$$f \longmapsto Ev_x(f) := f(x). \tag{15}$$

[9] We adopt this change (f instead of M_{met}) to follow Rosen's writings since 1959 as closely as possible. Without any doubt his aim in using f to refer to metabolism was to emphasize that metabolism was a mapping (i.e. a function).

[10] The clearest enunciation is found in Rosen (1972).

Intuitively, if X and Y bear some extra structure, the set $H(X, Y)$ could consist of all "structure-preserving" mappings from X to Y. For instance if $X = \mathbb{R}^2$ and $Y = \mathbb{R}$ then one choice for $H(X, Y)$ could be the set of all differentiable mappings from X to Y; another choice could be the set of polynomial mappings from X to Y.

Notice that as soon as we can find, for any pair of distinct points $x, x' \in X$, a function $f \in H(X, Y)$ such that $f(x) \neq f(x')$, the correspondence $x \mapsto Ev_x$ is one-to-one. This allows us to identify the element x with the mapping Ev_x, thus embedding the set X as a subset $\hat{X}$ of $H(X, Y)$.

Rosen's central result depends crucially on finding the special circumstances under which the evaluation mappings Ev_x may be one-to-one.[11] Evaluation mappings, in the many different contexts where they arise, are not normally one-to-one. For example, in the generic, non-specific case where X and Y are just sets and $H(X, Y)$ is the set $Map(X, Y)$ of *all* mappings from X to Y, we see that for any given x in X there are many mappings f that take a prescribed value y at x, unless X or Y are one-point sets. On the other hand, if the set $H(X, Y)$ is constrained to be a suitable proper subset of $Map(X, Y)$ there can be one-to-one evaluation maps Ev_x at many points x in X (see Fig. 5).

In the case where Ev_x happens to be one-to-one, it in fact defines an embedding of the set $H(X, Y)$ into Y, and so the following Remark holds.

Remark. If $Ev_x : H(X, Y) \to Y$ is invertible for some $x \in X$, then we have $|H(X, Y)| \leqslant |Y|$.[12]

A corollary is that for Ev_x to be one-to-one it is necessary that $H(X, Y)$ be a proper subset of $Map(X, Y)$.

Conjecture. *There are meaningful circumstances under which Ev_b may be invertible, i.e. one-to-one (mathematical version of Rosen's Central Result).*

Rosen's insight was to realize that, although evaluation mappings are not in general invertible (one-to-one), it is possible to envisage situations where they are, at least for some (generic) points x in X. In these hypothetical situations (of which, unfortunately, he never gave any examples) the assumption of invertibility does not introduce any inconsistencies. Instead, the invertibility of an evaluation map Ev_x suggests some sort of implicit mathematical structure on X and Y that is preserved by the mappings f in $H(X, Y)$, and admits x as some sort of "generator", so that every f in $H(X, Y)$ is "betrayed" by its value $f(x)$ at x. This corresponds to the idea that our mappings f are rather "rigid", as their behaviour is completely determined by the value they assign to the single point x (Fig. 5).

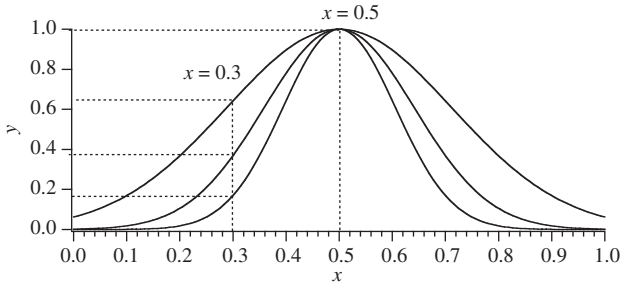

Fig. 5. Rigidity implied by invertible evaluations. Let $X = Y = [0, 1]$ and $H(X, Y)$ to be the infinite family of bell-shaped graphs ($y = f_\alpha(x) = \exp(-((x - 0.5)/\alpha)^2)$) sketched in the figure. Note that Ev_x is one-to-one for almost every x in $[0, 1]$ (for example at $x = 0.3$). The only exception is $x = 0.5$ because all curves collapse at this point.

Thus, Rosen's conjecture demands:

- Heavy restrictions on the type of allowed functions between X and Y. If $H(X, Y) = Map(X, Y)$, Rosen's conjecture is definitely false; it only holds under the very restricted condition $|H(X, Y)| \leqslant |Y|$.
- The elements $x \in X$, where evaluations are one-to-one, must have special properties.

5.2. β in (M, R) systems with organizational invariance

Rosen applied the previous result about invertibility to (M, R) systems with organizational invariance. His first step was to consider the functional relations of generic (M, R) systems (i.e. ones that are not necessarily organizationally invariant), which can be described by the following set of relations between maps:[13]

$$A \xrightarrow{f} B \xrightarrow{\Phi} H(A, B).$$

To apply the conjecture about invertibility, let us take $X = B$ and $Y = H(A, B)$. The construction $Ev_x : H(X, Y) \longrightarrow Y$ is transformed to $Ev_b : H(B, H(A, B)) \longrightarrow H(A, B)$.

This rather imposing formalism can be made a little more accessible by noticing that:

$H(A, B)$ represents the set $\mathscr{M}$ of possible metabolisms,

$H(B, H(A, B))$ represents the set $\mathscr{S}$ of possible selectors.

Then, Ev_b is a mapping that evaluates all possible selectors (i.e. all possible choices for Φ) at the element b, and $Ev_b \in Map(\mathscr{S}, \mathscr{M})$

$$Ev_b : \mathscr{S} \longrightarrow \mathscr{M},$$
$$\Phi \longmapsto Ev_b(\Phi) = \Phi(b). \tag{16}$$

[11] It is clear that Rosen demanded, not the invertibility of Ev_x, but that Ev_x should be one-to-one (injective). In fact to say "one-to-one" is equivalent to saying that Ev_x is invertible from X to $\hat{X}$. We maintain Rosen's term *invertibility* in most of the text.

[12] $|A|$ is the number of elements of the set A.

[13] This corresponds to the construction in Rosen (1972, p. 236).

The embedding

$$X \overset{Ev}{\hookrightarrow} Map(H(X, Y), Y)$$

becomes

$$B \overset{Ev}{\hookrightarrow} Map(H(B, H(A, B)), H(A, B))$$

or, in a more biological version,

$$B \overset{Ev}{\hookrightarrow} Map(\mathscr{S}, \mathscr{M}).$$

This apparently trivial result is heavy in consequences. It states that metabolic configurations (represented by $b \in B$), produced by a given metabolism, are equivalent (via the embedding construction) to Ev_b mappings. This statement is true for all (M, R) systems. But if at least one Ev_b is invertible (Conjecture) it means that there exists a mapping $Ev_b^{-1} \in Map(\mathscr{M}, \mathscr{S})$ such that

$$Ev_b^{-1}(f) = \Phi.$$

This mapping, which is none other than Rosen's β, is then given by (and may be identified with) an element $b \in B$. It follows that an organizationally invariant (M, R) system, represented by Rosen as

$$A \xrightarrow{f} B \xrightarrow{\Phi} H(A, B) \xrightarrow{\beta} H(B, H(A, B))$$

may be interpreted as two (M, R) systems acting in parallel:

$$(M, R)_{internal} : A \xrightarrow{f} B \xrightarrow{\Phi} \mathscr{M},$$

$$(M, R)_{external} : B \xrightarrow{\Phi} \mathscr{M} \xrightarrow{\beta} \mathscr{S}.$$

- In these two coupled (M, R) systems Φ is the replacement function for one and the metabolic function for the second.
- For this decomposition to hold, β must be the inverse of some evaluation map $Ev_b \in Map(\mathscr{S}, \mathscr{M})$.
- Invertible evaluations exist if the set $\mathscr{S}$ is small enough (hence $\mathscr{M}$ should also be small). To a first approximation $|B| = |\mathscr{M}| = |\mathscr{S}|$.
- If β exists, since it must be the inverse of some evaluation map $Ev_b \in Map(\mathscr{S}, \mathscr{M})$, for some $b \in B$, it may be identified with this $b \in B$, i.e. with a metabolic configuration.
- Rosen's formulation, which sometimes appears to indicate that $\mathscr{M} = Map(A, B)$ and $\mathscr{S} = Map(B, Map(A, B))$, is incorrect, because $\mathscr{M}$ and $\mathscr{S}$ are too big.

Rosen represented all these relations by Fig. 6.

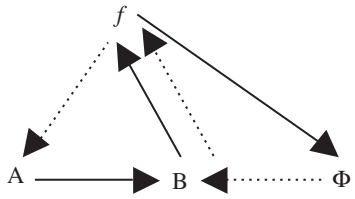

Fig. 6. This diagram represents Rosen's explanation of the closure of an (M, R) system. The broken arrows indicate that a function (located at the start of the arrow) uses a variable (at the destination) to produce a result. Each solid arrow indicates a transformation. The interesting aspect of this diagram is that every biological function (metabolism (f), replacement (Φ) or organizational invariance (the implied β which is equivalent to some element $b \in B$)) is entailed by another element in the diagram. No outside causality is needed. This is the basic bootstrapping property of a living system that justifies Rosen's statement that "Organisms are closed to efficient causes". A full discussion is in Chapter 10 of *Life Itself*, but *note* that Diagram 10.5C therein contains an error in defining the arrows.

6. Examples of (M, R) systems

A problem that every student of Rosen's works must confront is the astonishing lack of examples of the theoretical ideas in them. In this section, therefore, we introduce two examples of (M, R) systems. One is an arithmetical construction and the other uses a minimal, and ideal, metabolic network made of three reactions. Both examples serve to exemplify the notions of Φ and β. Neither is entirely satisfying; the arithmetical example is too remote from biological reality, and the metabolic one does not satisfy all of the requirements for an organizationally invariant system. Nonetheless, we consider these examples useful as steps towards a better understanding of Rosen's ideas.

6.1. An arithmetical example of a (M, R) system

Let $A = B = \mathbb{Z}_{12}$, the integers modulo 12 endowed with the operations of addition and multiplication modulo 12.[14] Note that some non-zero elements may have 0 as product (e.g. $3 \cdot 4 = 0 \bmod 12$), while others, called *invertible elements*, may be cancelled (e.g. $5 \cdot x = 5 \cdot y \Rightarrow x = y$).

We take $\mathscr{M} = H(A, B) = H(A, A)$ to consist only of the mappings f from A to B that preserve addition modulo 12, i.e. $f(x + y) = f(x) + f(y) \bmod 12$. This amounts to define $\mathscr{M}$ as $\mathscr{M} = \{f_c \mid c \in A\}$, where $f_c(x) = c \cdot x \bmod 12$. For example, the metabolism f_7 transforms metabolic state $a = 5$ into state $b = 11$, as $f_7(5) = 7 \cdot 5 = 11 \bmod 12$. We remark that the set $\mathscr{M}$ is naturally endowed with the operation of pointwise addition:

$f_c + f_d : x \mapsto f_c(x) + f_d(x) = cx + dx = (c + d)x = f_{c+d}(x)$. Notice that $f_c + f_d$ is just f_{c+d}.

Thus, using this simple algebraic framework, an (M, R) system is represented by

$$A \xrightarrow{f_c} A \xrightarrow{\Phi} H(A, A) = \mathscr{M}.$$

[14] $\mathbb{Z}_{12} \coloneqq \mathbb{Z}/12\mathbb{Z} = \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11\}$ and, for instance, $8 + 7 = 15 = 3 \bmod 12$ and $7 \cdot 4 = 28 = (2 \cdot 12 + 4) = 4 \bmod 12$.

We now need to specify the set $\mathscr{S} = H(A, \mathscr{M})$ where the selectors Φ are to be chosen. Since we have addition modulo 12 on $A = \mathbb{Z}_{12}$, and an analogous addition modulo 12 on $\mathscr{M}$, we will ask that the mappings Φ in $\mathscr{S}$ transform addition modulo 12 in A into addition modulo 12 in $\mathscr{M}$, according to

$$\Phi(c + d) = \Phi(c) + \Phi(d).$$

It follows that $\mathscr{S}$ consists of the elements Φ_k ($k \in A$) given by[15]

$$\Phi_k(1) = f_k,$$
$$\Phi_k(c) = c\Phi_k(1) = cf_k = f_{ck}.$$

Thus, the basic (M, R) system (where $f = f_c$) can now be written as

$$A \xrightarrow{f_c} A \xrightarrow{\Phi_k} \mathscr{M}$$

with the following equations:

$$\begin{aligned} b &= f_c(a) \quad \text{(Metabolism)},\\ \Phi_k(b) &= f_c \quad \text{(Replacement)}. \end{aligned} \tag{17}$$

These translate to the following equations:

$$\begin{aligned} b &= f_c(a) \quad \text{i.e. } b = ca \bmod 12,\\ f_{bk} &= f_c \quad \text{i.e. } bk = c \bmod 12 \end{aligned}$$

which, as $b = ca$, reduce to $cak = c \bmod 12$, which is equivalent to $ak = 1 \bmod 12$ if c is invertible. Thus, the equation $\Phi_k(f_c(a)) = f_c$, for k, may have a unique solution, several solutions or no solution at all, depending on the invertibility of c (which defines f_c) and a. We illustrate these properties with some examples:

Case 1: (*a invertible*, *c non-invertible*, *two solutions*).

Let us choose $a = 5$, $f = f_2$. Then $b = f_2(5) = 2 \cdot 5 = 10 \bmod 12$. Let us look for $\Phi = \Phi_k$ such that $\Phi_k(10) = f_2$. This is equivalent to solving the equation $10k = 2$ in A, which has solutions $k = 5$ and $11 \bmod 12$. We see that in this case there are two mappings Φ, namely Φ_5 and Φ_{11}, such that $\Phi(b) = f$.

Case 2: (*a invertible*, *c non-invertible*, *three solutions*).

Let us choose $a = 5$, $f = f_3$. Then $b = f_3(5) = 3 \cdot 5 = 15 = 3 \bmod 12$. Let us look for $\Phi = \Phi_k$ such that $\Phi_k(3) = f_3$. This is equivalent to solving the equation $3k = 3$ in A, which has solutions $k = 1, 5$ and $9 \bmod 12$. So in this case there are three mappings Φ, namely Φ_1, Φ_5 and Φ_9, such that $\Phi(b) = f$. Note that the difference with the previous example arises from the fact that the equation $3k = 0$ has three solutions (0, 4 and 8) while $2k = 0$ has only two (0 and 6).

Case 3: (*a non-invertible*, *c invertible*, *no solution*).

Let us choose $a = 4$, $f = f_5$. Then $b = f_5(4) = 5 \cdot 4 = 20 = 8 \bmod 12$. Let us look for $\Phi = \Phi_k$ such that $\Phi_k(8) = f_5$. This is equivalent to the equation $8k = 5$, which has no solution in A.

Case 4: (*a invertible*, *c invertible*, *one solution*).

Let us choose $a = 5$, $f = f_7$. Then $b = f_7(5) = 7 \cdot 5 = 35 = 11 \bmod 12$. We need to find Φ_k such that $\Phi_k(11) = f_7$. But this is equivalent to solving the equation $11k = 7$ in A, which has the unique solution $k = 5 \bmod 12$. We have therefore, only one Φ, namely Φ_5 such that $\Phi(b) = f$. We now see that, for $b = f_c(a) = 11$, which forces a and c to be invertible, we can find a unique Φ such that $\Phi(b) = f_c$, namely $\Phi = \Phi_k$ where k is the *only* solution of the equation $11k = c$ in A, i.e. $k = 11c$. So for this particular choice of b the invertibility of the evaluation at b is fulfilled, i.e. the core of Rosen's central result is satisfied. Following Rosen's terminology, we have then $\beta(f_c) = \Phi_{11c}$, for all $f_c \in \mathscr{M}$.

This example suggests that, in order to have an organizationally invariant (M, R) system it is necessary to restrict both the elements b and the type of allowed mappings f (metabolisms) and Φ (selectors). The example also disposes of the suggestion by Landauer and Bellman (2002) that the invertibility required by Rosen is not possible, even in a formal sense.

6.2. A metabolic example of an (M, R) system

An example with a more biological flavour, built on one suggested by Morán et al. (1996), can be constructed from simplified rules representing idealized metabolisms. Consider the following three reactions, without specifying the nature of the catalysts M_1, M_2, and M_3, which represent a minimal metabolism:

$$s + t \xrightarrow{M_1} st,$$
$$s + u \xrightarrow{M_2} su,$$
$$st + u \xrightarrow{M_3} stu.$$

These three equations specify a particular instance of a metabolism M between the sets $A = \{a\} = \{(s,t),(s,u),(st,u)\}$ and $B = \{b\} = \{(st, su, stu)\}$. Here the set $\mathscr{M}$ is simply *one* transformation f such that $f((s,t),(s,u),(st,u)) = (st, su, stu)$, i.e. $f(a) = b$.

To specify the corresponding R part, the subsystem of metabolic reactions producing each R_i must be specified. In this simplified network this specification is simply to identify one of the outputs $\{st, su, stu\}$ with one of the M_i (thus we are specifying the production mechanism by which a given R_i is continuously generated). A great number of assignments are possible, in total 3^3, but

[15] Note that in this example sets $\mathscr{M}$ and $\mathscr{S}$ contain the same number of elements (12) and this number is not bigger than $|A| = 12$. We claim that this property is a general characteristic of (M, R) systems with organizational invariance.

this number[16] is decreased substantially by excluding "self-catalytic" assignments such as $M_1 = st$, or $M_3 = stu$, in which the product of a reaction catalyses the same reactions that produce it. Although the point is arguable, and others may arrive at a different conclusion, we think that autocatalysis of this kind should be avoided in the theory of (M, R) systems as we are seeking systems in which the circularity is a property of the global connectivity in the entire network and not a property of a single reaction. This restriction requires, for example, the only valid assignment for $st + u \xrightarrow{M_3} stu$ to be $M_3 = su$ (as su is neither a substrate nor a product of reaction 3). This kind of argument decreases the initial 27 possibilities to the following four valid assignments for Φ:

$$\Phi_1 : (M_1, M_2, M_3) \longrightarrow (stu, stu, su),$$
$$\Phi_2 : (M_1, M_2, M_3) \longrightarrow (stu, st, su),$$
$$\Phi_3 : (M_1, M_2, M_3) \longrightarrow (su, stu, su),$$
$$\Phi_4 : (M_1, M_2, M_3) \longrightarrow (su, st, su).$$

Each of these selectors generates a different (M, R) system, where the M part is similar:

$$(M, R_1) = \begin{array}{c} s + t \xrightarrow{stu} st \\ s + u \xrightarrow{stu} su \\ st + u \xrightarrow{su} stu \end{array}$$

$$(M, R_2) = \begin{array}{c} s + t \xrightarrow{stu} st \\ s + u \xrightarrow{st} su \\ st + u \xrightarrow{su} stu \end{array}$$

$$(M, R_3) = \begin{array}{c} s + t \xrightarrow{su} st \\ s + u \xrightarrow{stu} su \\ st + u \xrightarrow{su} stu \end{array}$$

$$(M, R_4) = \begin{array}{c} s + t \xrightarrow{su} st \\ s + u \xrightarrow{st} su \\ st + u \xrightarrow{su} stu. \end{array}$$

Among these Φ_4 is special, as the third reaction of (M, R_4) ($st + u \xrightarrow{su} stu$) does not participate in the network because stu is neither the substrate nor the catalyst of another reaction in this small network. We therefore discard Φ_4. Thus, from the 27 choices for Φ that are theoretically compatible with this simple metabolism f we have discarded 24, leaving only three as valid assignments, and so the set of selectors is reduced to $\mathscr{S} = \{\Phi_1, \Phi_2, \Phi_3\}$.

The procedure outlined here, starting with the information provided by f and serving to define the set of possible selectors Φ is an embodiment of the function β, which turns out here to be a "multivalued function": $\beta(f) = \{\Phi_1, \Phi_2, \Phi_3\}$.

The fact that $\beta(f)$ is not single-valued (as any honest function should be) shows that the condition of invertibility, which is the defining property of (M, R) systems with organizational invariance, fails for this simple metabolic network. Thus, although this metabolic network is an (M, R) system, and also an autocatalytic network (Kauffman, 1993), it cannot be construed as an organizationally invariant (M, R) system because the rule for assigning Φ starting from f gives more than one result. This example is also interesting as it shows that an autocatalytic set (such as the one discussed in this section) is not necessarily an (M, R) system with organizational invariance. The two ideas, although related, are different in a fundamental way.

7. Infinite regress, self-reference and endomorphisms

As we have stated, Rosen's central result was the escape route he devised to avoid the infinite chain of repairers of repairers of repairers that is implied by the core mechanism of (M, R) systems. Surprisingly, the three-step chain of mappings and sets that constitutes the framework of the central result suggests a construction that generates mathematical objects that are solutions to the puzzling-looking equation $f(f) = f$.

As previously shown, the algebraic construction of an organizationally invariant (M, R) system is given by

$$A \xrightarrow{f} B \xrightarrow{\Phi} H(A, B) \xrightarrow{\beta} H(B, H(A, B))$$

with the following crucial ingredients:

$$f(a) = b \quad \text{with } f \in H(A, B) = C_0,$$
$$\Phi(b) = f \quad \text{with } \Phi \in H(B, H(A, B)) = C_1,$$
$$\beta(f) = \Phi \quad \text{with } \beta \in H(H(A, B)), H(B, H(A, B)) = C_2,$$

β being the inverse of the (putatively invertible) "evaluation at b" mapping Ev_b.

This construction has two peculiarities: f plays, in the different steps, the roles of function, variable and result; the sets C_i are built according to a recursive rule, namely $C_n = H(C_{n-2}, C_{n-1})$. If we intend to set up an infinite

[16] This may not appear so huge, but bear in mind that the model, with just three steps, is very small. A biologically more realistic example with, say, 3000 steps, would have 3000^{3000} possible assignments, a truly gigantic number (much bigger than $10^{10\,000}$).

extension of this chain, notation must be consolidated as $C_0 = A$, $C_1 = B, C_2 = H(C_0, C_1)$, and in general

$$C_n = H(C_{n-2}, C_{n-1}).$$

With respect to the mappings involved, we define $\Phi_0 = f, \Phi_1 = \Phi, \Phi_2 = \beta$. For the transformed elements we define $c_0 = a$, $c_1 = b$, $c_2 = f$.

Thus, the three building relations of an (M, R) system become

$$c_1 = \Phi_0(c_0) \quad \text{(equivalent to } f(a) = b\text{)},$$
$$c_2 = \Phi_1(c_1) = \Phi_0 \quad \text{(equivalent to } \Phi(b) = f\text{)},$$
$$c_3 = \Phi_2(c_2) = \Phi_1 \quad \text{(equivalent to } \beta(f) = \Phi\text{)}.$$

Under mild assumptions on the sets $H(C_{n-2}, C_{n-1})$ of structure-preserving mappings from C_{n-2} to C_{n-1}, we may extend these recurrence relations as

$$c_{n+1} = \Phi_n(c_n) = \Phi_{n-1}, \quad \text{so that } \Phi_n(\Phi_{n-2}) = \Phi_{n-1}.$$

Thus, a natural extension of the three steps of Rosen's central result is

$$C_0 \xrightarrow{\Phi_0} C_1 \xrightarrow{\Phi_1} C_2 \xrightarrow{\Phi_2} \cdots C_n \xrightarrow{\Phi_n} C_{n+1} = H(C_{n-1}, C_n),$$
$$c_o \longmapsto c_1 \longmapsto c_2 \longmapsto \cdots c_n \longmapsto c_{n+1} = \Phi_n(c_n) = \Phi_{n-1},$$

where Φ_2 is no longer assumed to be the inverse of the "evaluation at c_1" map Ev_b.

If we assume that these sequences converge to limits Φ_∞ and C_∞, which is indeed always the case, at least formally (Soto-Andrade and Varela, 1984), then passing to the limit in the above equations, we have that

$$C_n = H(C_{n-2}, C_{n-1}) \xrightarrow{n\to\infty} C_\infty = H(C_\infty, C_\infty),$$
$$\Phi_{n-1} = \Phi_n(\Phi_{n-2}) \xrightarrow{n\to\infty} \Phi_\infty(\Phi_\infty) = \Phi_\infty.$$

These two limits exhibit two fundamental properties:

- $C_\infty = H(C_\infty, C_\infty)$, which means that C_∞ is a reflexive domain, a structured set that is equal to the set of its endomorphisms;
- $\Phi_\infty(\Phi_\infty) = \Phi_\infty$ and so Φ_∞ is a self-referential object, i.e. a solution to $f(f) = f$.

Thus, Rosen's basic construction behind his central result could be extended to an infinite chain to define limiting objects with interesting properties for the study of circularity. Surprisingly, it seems that Rosen, starting from a biological insight, was the first to notice how to define this type of mathematical object. A question not addressed here is to find the relation between Φ_∞ and Rosen's β (the inverse to the evaluation at b map Ev_b), as our infinite chain can be constructed without any reference to the possibility or otherwise of inverting the Ev_{c_i} maps (unpublished work).

8. Discussion

As stated in the introduction, Rosen's work is often cryptic and unknown, and the main objective of this paper has been to clarify some central aspects of his ideas. Without such clarification his work would suffer 40 more years of obscurity. We hope that the clarification achieved here will encourage other researchers to read his publications critically and to apply them to current problems in metabolic analysis and the origin of living systems.

Apart from clarifying Rosen's theory, in itself a major task, our main theoretical contribution has been to show that the central result holds when the sets $\mathscr{M}$ and $\mathscr{S}$ are restricted. This is an important result, as it shows that, as well as the purely logical constraints hinted at by Rosen, it is necessary to add to the model of (M, R) systems ad hoc restrictions on the types of metabolisms and selectors that are allowed. The generic statement $A \xrightarrow{f} B$, with $f \in Map(A, B)$ is simply too general. Metabolism should be more complex than a single arrow connecting two sets. A similar argument applies of course to $\Phi \in Map(B, Map(A, B))$. We have discovered that $|B| = |\mathscr{M}| = |\mathscr{S}|$; thus the amount of selectors ($|\mathscr{S}|$) must be of the order of the elements of molecules in the metabolism ($|B|$). These restrictions may take many forms, from forbidding all autocatalytic reactions, to the obvious necessity that in every real biochemical reaction the law of conservation of mass must hold.

The difficulty for understanding Rosen's ideas arises from several causes. First, his ideas may appear on first reading to be little more than an application of graph theory to metabolic networks. This misidentification may then wrongly suggest that (M, R) systems are just a 1950s version of current topics in metabolic analysis (Hofmeyr and Cornish-Bowden, 2000). Any such interpretation misses the essential fact that (M, R) systems embody a unique attempt to prove that metabolic networks, which constitute the foundation of living systems, must satisfy certain logical regularities that go beyond stoichiometric or thermodynamic constraints. These logical regularities arise from the circular nature of biological organization, which can be summarized by saying that the three-step chain of mappings and sets that constitutes the framework of the central result suggests a construction that generates mathematical objects that are solutions to the puzzling-looking equation $f(f) = f$.

Second, Rosen's main idea is out of the mainstream of current theoretical models, as it is an algebraic theory totally uncoupled from the usual language of differential equations. His central result refers to something most biologists will find extremely esoteric: an attempt to prove (from purely logical grounds) the circularity or closure of metabolic networks. Furthermore, the mathematical concept used to introduce this notion, the

invertibility of evaluation mappings, is unusual enough to make it very difficult to explain the context of the result even to mathematicians. Rosen's use of category theory erects a further barrier to comprehension for many readers.

Third, Rosen, for reasons that are difficult to understand, never specified the mathematical context where his insight could hold true, and never gave an example, either mathematical or biological. The very few people who have tried to follow his scientific path have quoted his conclusions, but have rarely explained them and have almost never provided examples.

As may be surmised, we have adopted the point of view that Rosen had a powerful insight on the nature of metabolic networks and the necessary (but otherwise ignored) requirement of circularity. Our analysis differs radically from a recent analysis of Rosen's ideas which states that "unfortunately, the mathematics [of his analysis] is incorrect, and the assertions remain unproven (and some of them are simply false)" (Landauer and Bellman, 2002). Landauer and Bellman are mistaken, and their conclusions are incorrect, as they are based only on the summary of his ideas that Rosen provided in *Life Itself* (Chapter 10) and not in the study of the notion of invertibility of β and the essential papers of 1958, 1959 and 1972. Our arithmetical (M,R) example explicitly invalidates Landauer and Bellman's critique, as it has a mathematical structure with all the properties that the central result demands. Our analysis also differs from the extensive work of Casti (1988), in which he analysed a subset of (M,R) systems with the techniques of linear analysis (and systems identification). In his analysis the true nature of β again appears to be problematical, to the point that Casti has to introduce its existence as a specific postulate (Casti, 1988, p. 116). Our analysis goes further than his, as we tackled the nature of β and found a condition, $|B| = |\mathscr{M}| = |\mathscr{S}|$, under which a non-trivial β can exist for a generic (M,R) system, not just the special case of linear behaviour.

However, Rosen's insight is far from being workable and ready to apply to current network analysis without major efforts, both to clarify the circumstances in which his central theorem applies, and to explain its meaning in biological terms. An intriguing possibility could be to combine his analysis with the notion of autopoietic systems, another theory that posits metabolic closure as the core of biological organization (Maturana and Varela, 1980; Varela et al., 1974; Letelier et al., 2003). This paper is intended to make a step in the proper direction, as we have isolated from Rosen's extensive work what we think is its main point, and we have advanced, albeit not to a stage that will satisfy everyone, in clarifying concepts like, $f : A \to B$ (metabolism viewed as a mapping), $\Phi : B \to H(A,B)$ ("repair", in Rosen's sense, as replacement) and $\beta : H(A,B) \to H(B,H(A,B))$ (organizational invariance). We have explained the mathematical insight behind the idea of organizational invariance as embodied in the operator β, a crucial concept that essentially acts as a *generator* of the complete formal structure of an (M,R) system. In effect it is possible to reformulate the definition of an organizationally invariant (M,R) system as the kind of system where for some b the equation $\Phi(b) = f$ has exactly one solution Φ, for any given f, giving rise to the operator β, which sends any f to its associated Φ and then implicitly giving the structure of the whole system. The explicit construction of β was referred to by Rosen as *the realization problem* (Rosen, 2000, p. 262) and he conceded that it was difficult at the level of the theory, or at the level of a physical model, to construct a metabolic network that would embody the notion of β.

Although we have not as yet been as successful in providing a satisfactory metabolic example of a minimal (M,R) system with organizational invariance, we believe that our analysis of why our (M,R) system of three reactions failed to be organizationally invariant is nonetheless informative and illuminates Rosen's main ideas.

As must be clear by now, the core of Rosen's view of metabolic organization cannot be summarized or explained in few pages. Furthermore, here we have not touched on his contributions to the theory of complex systems (Rosen, 1985, 1991). His framework is unique in the sense that he approaches biological organization in an intrinsically non-reductionist manner. In fact he never talks about physical particles (genes, enzymes, etc.) at all, but instead considers system-wide functions (metabolism, organizational invariance). In this way his approach is a step towards understanding components in relation to whole systems (Cornish-Bowden et al., 2004), moving away from the reductionist tradition of treating whole systems as little more than the sum of their parts.

Acknowledgments

We thank Diane Greenstein for her editorial skills. This work was supported by Fondecyt 1030371 (JCL), Fondecyt 1040444 (JSA) and the CNRS (AC-B, MLC). We thank Jorge Mpodozis and Gonzalo Marín for introducing a healthy degree of scepticism into our discussions, and also Randall Thomas for considerable help in locating some of Rosen's publications.

References

Barbieri, J., 2005. Life is "artifact-making". J. Biosemiotics 1, 107–134.

Cárdenas, M.L., Cornish-Bowden, A., Ureta, T., 1998. Evolution and regulatory role of the hexokinases. Biochim. Biophys. Acta. 1401, 242–264.

Casti, J., 1988. The theory of metabolism-repair systems. Appl. Math. Comput. 28, 113–115.

Casti, J., 2002. "Biologizing" control theory: how to make a control system come alive. Complexity 7, 10–13.

Cornish-Bowden, A., Cárdenas, M.L., Letelier, J.-C., Soto-Andrade, J., Guíñez-Abarzúa, F., 2004. Understanding the parts in terms of the whole. Biol. Cell 96, 713–717.

Hofmeyr, J.-H.S., Cornish-Bowden, A., 2000. Regulating the cellular economy of supply and demand. FEBS Lett. 476, 47–51.

Joslyn, C., 1993. Review of Life Itself. Int. J. Gen. Systems 21, 394–402.

Kauffman, S., 1993. The Origins of Order: Self-organization and Selection in Evolution. Oxford University Press, Oxford.

Landauer, C., Bellman, K., 2002. Theoretical biology: organisms and mechanisms. AIP Conf. Proc. 627, 59–70.

Letelier, J., Marín, G., Mpodozis, J., 2003. Autopoietic and (M, R) systems. J. Theor. Biol. 222, 261–272.

Letelier, J.-C., Soto-Andrade, J., Guíñez Abarzúa, F., Cornish-Bowden, A., Cárdenas, M.L., 2004. Metabolic closure in (M, R) systems. In: Pollack, J., Bedau, M., Husbands, P., Ikegami, T., Watson, R.A. (Eds.), Artificial Life, vol. IX. MIT Press, Cambridge, MA, pp. 450–455.

Maturana, H., Varela, F., 1980. Autopoiesis and Cognition: the Realization of the Living. Reidel, Dordrecht.

Mikulecky, D., 2001. Robert Rosen (1934–1998): a snapshot of biology's Newton. Comput. Chem. 25, 317–327.

Morán, F., Moreno, A., Minch, E., Montero, F., 1996. Further steps towards a realistic description of the essence of life. In: Langton, C.G., Shimohara, K. (Eds.), Artificial Life, vol. V. MIT Press, Cambridge, MA, pp. 255–263.

Nomura, T., 2002. Formal description of autopoiesis for analytical models of life and social systems. In: Standish, R., Bedau, M., Abbass, H. (Eds.), Artificial Life, vol. VIII. MIT Press, Cambridge, MA, pp. 15–18.

Pierce, B., 1991. Basic Category Theory for Computer Scientists. MIT Press, Cambridge, MA.

Rosen, R., 1958a. A relational theory of biological systems. Bull. Math. Biophys. 20, 245–341.

Rosen, R., 1958b. The representation of biological systems from standpoint of the theory of categories. Bull. Math. Biophys. 20, 317–341.

Rosen, R., 1959. A relational theory of biological systems II. Bull. Math. Biophys. 21, 109–128.

Rosen, R., 1966. A note on replication in (M, R) systems. Bull. Math. Biophys. 28, 149–151.

Rosen, R., 1967. Further comments on replication in (M, R) systems. Bull. Math. Biophys. 29, 91–94.

Rosen, R., 1971. Some realizations of (M, R) systems and their interpretation. Bull. Math. Biophys. 33, 303–319.

Rosen, R., 1972. Some relational cell models: the metabolism-repair system. In: Rosen, R. (Ed.), Foundations of Mathematical Biology. Academic Press, New York.

Rosen, R., 1985. Anticipatory Systems. Pergamon Press, Oxford.

Rosen, R., 1991. Life Itself. Columbia University Press, New York.

Rosen, R., 2000. Essays on Life Itself. Columbia University Press, New York.

Soto-Andrade, J., Varela, F., 1984. Self-reference and fixed points: a discussion and an extension of Lawvere's theorem. Acta Appl. Math. 2, 1–19.

Varela, F., Maturana, H., Uribe, R., 1974. Autopoiesis: the organization of living systems, its characterization and a model. BioSystems 5, 187–196.

Wolkenhauer, O., 2001. Systems biology: the reincarnation of systems theory applied to biology? Briefings Bioinf. 2, 258–270.

Wolkenhauer, O., 2002. Mathematical modelling in the post-genome era: understanding genome expression and regulation—a systems theoretic approach. BioSystems 65, 1–18.