THEORETICAL BIOLOGY AND COMPLEXITY

Three Essays on the Natural Philosophy of Complex Systems

Edited by

ROBERT ROSEN

Contributors

A. H. Louie
I. W. Richardson
Robert Rosen

Theoretical Biology and Complexity

Three Essays on the Natural Philosophy of Complex Systems

Edited by

Robert Rosen

Department of Physiology and Biophysics
Dalhousie University
Halifax, Nova Scotia

1985

ACADEMIC PRESS, INC.

(Harcourt Brace Jovanovich, Publishers)

Orlando San Diego New York London
Toronto Montreal Sydney Tokyo

ACADEMIC PRESS, INC.
Orlando, Florida 32887

United Kingdom Edition published by
ACADEMIC PRESS INC. (LONDON) LTD.
24–28 Oval Road, London NW1 7DX

Library of Congress Cataloging in Publication Data
Main entry under title:

Theoretical biology and complexity.

Contents: The dynamics and energetics of complex real systems / I. W. Richardson -- Categorical system theory / Aloisius Ho-Yin Louie -- Organisms as causal systems which are not mechanisms / Robert Rosen.
1. Biological systems--Addresses, essays, lectures. 2. Biological systems--Measurement--Addresses, essays, lectures. 3. System theory--Addresses, essays, lectures. I. Richardson, I. W., DATE. Dynamics and energetics of real systems. 1985. II. Louie, Aloisius Ho-Yin. Categorical system theory. 1985. III. Rosen, Robert. Organisms as causal systems which are not mechanisms. 1985.
QH311.T48 1985 574 84-21568
ISBN 0-12-597280-6 (alk. paper)

PRINTED IN THE UNITED STATES OF AMERICA

85 86 87 88 9 8 7 6 5 4 3 2 1

List of Contributors

Numbers in parentheses indicate the pages on which the authors' contributions begin.

*A. H. Louie** (69), Department of Physiology and Biophysics, Dalhousie University, Halifax, Nova Scotia B3H 4H7, Canada

I. W. Richardson† (1), Department of Physiology and Biophysics, Dalhousie University, Halifax, Nova Scotia B3H 4H7, Canada

Robert Rosen (165), Department of Physiology and Biophysics, Dalhousie University, Halifax, Nova Scotia B3H 4H7, Canada

* Present address: 86 Dagmar Avenue, Vanier, Ontario K1L 5T4, Canada. No longer affiliated with Dalhousie University.

† Present address: 908 Middle Avenue, Menlo Park, California 94025. No longer affiliated with Dalhousie University.

Preface

This volume is made up of three short essays, each separately conceived and written, each with distinct thrusts and emphases, but nevertheless closely related in substance and spirit. The purpose of this brief introduction is to describe some of these relations, which are both personal and scientific.

Let us turn first to the scientific threads that relate the three contributions. Although very different in emphasis and in thrust, they all spring from a common concern: to grasp and comprehend the material basis of living systems. I believe that we each began with a conviction that contemporary physics already contained the necessary universals with which to cope with the phenomena of life and that therefore only a clever rearrangement and redeployment of these universals would suffice to bring them to bear effectively on biology. I believe that we each came separately, and with great reluctance, to admit the possibility that this conviction might not be true, and hence that a true theory of the organism required new physics and new epistemology. And again separately, we realized that the measurement process, which lies at the very heart of every mode of system description, provides perhaps the only safe and fundamental point of departure for building a comprehensive theory, not only of organisms, but of natural systems in general. This premise is the primary thread that binds the essays in this volume together. As Dr. Richardson likes to say, such an approach restores to our fragmented sciences the kind of integration and unity they possessed in an earlier time, when scientists regarded themselves as natural philosophers.

In his essay, Dr. Richardson casts his analysis of the measurement process into an elegant dualism relating modes of description and explores the consequences of this really remarkable dualism for what we may call classical physics. But the dualism he develops pertains to any kind of description; rich as it is in the familiar contexts of classical physics, it becomes still richer when treated as a universal principle and brought to bear on organisms.

Dr. Louie explores the deeper consequences of representing the properties of natural systems through states built up out of observable quantities and the dynamics that such systems impose on each other through interactions. The natural mathematical universe for exploring these consequences is category theory, a mathematical discipline that is in a formal sense the general theory of modelling and that in this application provides both the mathematical tool and the best example for exploring relationships between systems. His work provides the natural bridge between Dr. Richardson's essay and my own; he explicitly describes a number of these relationships in the course of his development.

In my own essay, a rather radical viewpoint is adopted, motivated from several different considerations. One of these is my continuing involvement with relational biology, which is described in some detail. Another was my concern with systems that can anticipate, more specifically, ones that contain internal predictive models of themselves and/or their environments, whose predictions can be utilized to modify and control present activities. Still another is an attempt to construct a dictionary relating the language of physics (forces, potentials, fields) with the informational language (code, program, computation) so characteristic of biology, by using the concept of stability as Rosetta Stone. Consideration of these problems, separately and together, led ultimately to the suggestion that our traditional modes of system representation, involving fixed sets of states together with imposed dynamical laws, strictly pertains only to an extremely limited class of systems (which I call simple systems or mechanisms). Systems not in this class I call complex, and these can only be in some sense approximated, locally and temporally, by simple ones. Such a radical alteration of viewpoint leads to a large number of concrete, practical consequences, some of which are described in the essay.

All of these essays are the products of their authors' association with the Biomathematics Program of the Department of Physiology and Biophysics at Dalhousie University in Halifax, Nova Scotia. This program has become known as the Red House because it is housed in quarters that are, in fact, red. The creation of this program was largely the work of one of the authors, I. W. Richardson. Through his initiative, effort, and energy, he was able almost single-handedly to carve out, for a perhaps brief but precious time, what I regard as one of the most innovative and fruitful programs for research and teaching in theoretical biology in the world.

I became associated with the Red House in 1975 when, again largely through Dr. Richardson's efforts, I was offered a Killam Professorship at Dalhousie University. I found that Dr. Richardson had created an ideal atmosphere, an ambience that almost compelled those who experienced it to think, study, learn, and write beyond the edge of the known. I had experienced such ambiences before, at Chicago under Rashevsky and at Buffalo under J. F. Danielli. There are others, but they are few.

The Red House program has over the course of time attracted some extraordinary students. One of these is Dr. Aloisius Louie, whose contribution to this volume is essentially his doctoral dissertation. I would normally strongly discourage any student, however bright, from undertaking a dissertation involving a major epistemological component; dissertations are risky enough without that. But Dr. Louie is a very special case; he thrives on such risks and languishes without them. Indeed, it will be seen from the definitive discussion he presents of the category-theoretic roots of system analysis that he has been from the outset a colleague rather than a student.

It is thus with a certain sense of pride that we collectively offer this volume, not only for what novel scientific material it contains, but also as a specimen of the output of the Red House program and as a symbol for what such programs can accomplish. This last is important, for it is precisely the scientific strengths of such programs that also make them vulnerable, in constant threat of engulfment by the sands of the vast academic deserts which surround them.

We hope that these essays will provide the reader with some food for thought and will convey some of the electricity and excitement, as well as the practical import, of doing theory.

2

Categorical System Theory

A. H. Louie[†]

Department of Physiology and Biophysics
Dalhousie University
Halifax, Nova Scotia

[†] Present address: 86 Dagmar Ave., Vanier, Ontario. No longer affiliated with Dalhousie University.

ISBN 0-12-597280-6

LIST OF SPECIAL SYMBOLS

The symbols listed below are followed by a brief statement of their meaning. The standard set-theoretic symbols are not included here.

$\mathbf{C}$	General category
$\mathbf{C}(A, B)$	Hom set
$f: A \to B$	Morphism
$g \circ f$	Composition of morphisms
dom f	Domain
1_A	Identity morphism
f^{-1}	Inverse morphism
$A \cong B$	Isomorphic objects
Ens	Category of sets
B^A	$= \mathbf{Ens}(A, B)$
Gp	Category of groups
Top	Category of topological spaces
Mon	Category of monoids
Ab	Category of abelian groups
$F: \mathbf{A} \to \mathbf{B}$	Functor
$\mathbf{C}^{op}$	Dual category
f^{op}	Opposite morphism
h^A	Covariant hom-functor
h_A	Contravariant hom-functor
Cat	Category of (small) categories
$I_{\mathbf{A}}$	Identity functor
$F^{-1}: \mathbf{B} \to \mathbf{A}$	Inverse functor
$\alpha: F \to G$	Natural transformation
$\mathbf{B}^{\mathbf{A}}$	Functor category
ΠA_i	Product
$\pi_i: A \to A_i$	Projection
Σ	Statement
Σ^{op}	Dual statement
$\amalg A_i$	Coproduct
$\iota_j: A_j \to A$	Injection
S	Set of states
$f, g, h, \ldots$	Observables
$\mathbb{R}$	Real numbers
R_f	Equivalence relation induced by f
$f \sim g$	iff $R_f = R_g$, algebraic equivalence
$f \leq g$	iff $R_f \subset R_g$, refinement
$\mathbf{S}$	Category of formal systems
(S, F)	Formal system
F	Set of observables
$0.$	Constant observable
$\phi, \psi, \chi, \ldots$	$\mathbf{S}$-morphisms
R_G	$= \cap R_f: f \in G$
eq(f, g)	Equalizer
coeq(f, g)	Coequalizer
$f: A \rightarrowtail B$	Monomorphism
χ_B	Characteristic function
$f: A \twoheadrightarrow B$	Epimorphism
$(S_1, F_1) \leq (S_2, F_2)$	Partial order in $\mathbf{S}$
M	Meter
$\mathscr{A}(S)$	Group of automorphisms on S
$\mathbf{D}$	Category of dynamical systems
T	Dynamics
t	Time
$a_x[T], b_x[T]$	Bounds of T at x
$\bar{\mathbb{R}}$	$= [-\infty, +\infty]$, extended real numbers
I_S	Trivial dynamics
y_x	Solutions of a dynamics
$y_x(a_x, b_x)$	Trajectories of a dynamics
T_t	Translations of a dynamics
(S, D)	Dynamical system
D	Set of dynamics
$\phi, \psi, \chi, \ldots$	$\mathbf{D}$-morphisms
$\phi \times 1_{\mathbb{R}}$	$:(x, t) \mapsto (\phi x, t)$
$\mathbb{N}$	Natural numbers
$(\mathscr{E}, \mathscr{M})$	Image factorization system
$\mathscr{R}(S)$	Lattice of equivalence relations on S
R_T	Conjugate equivalence relation
L_T	Conjugate lattice

$\mathbb{Z}$	Integers	S^a	Activities of S
$\langle T \rangle$	$= \{T^n : n \in \mathbb{Z}\}$, discrete dynamics	S^p	Products of S
τ	Topology	S_1, S_2, S_3	Specialized subsets of S
d_f	f-induced pseudometric	t_i	Survival time of e_i
$E: N \to S$	Modelling relation	S_R	Product-induced relations
$\mathbf{N}$	Category of natural systems	$\mathbf{O}$	Organismic category
(S, F, D)	Natural system	V	Vector (inner product) space
Φ, Ψ	Linkage relations	V^*	Dual vector space
$(S_1, F_1, D_1) \leq (S_2, F_2, D_2)$	Partial order in $\mathbf{N}$	$\mathbf{F}, \mathbf{F}_i$	Contravariant vectors
		$\mathbf{a}, \mathbf{a}^i$	Covariant vectors
		$\mathbf{R} = \mathbf{aF}$	Dyad
$\mathbf{L}$	Organism	$T_1^1(V)$	Tensor space of type (1, 1) over V
L_i	$\mathbf{L}$-objects		
ϕ_{ji}	$\mathbf{L}$-morphisms	$\mathbf{R} = \Sigma_i \mathbf{a}^i \mathbf{F}_i$	Dyadic (= response tensor)
$\mathbf{K}$	Structured category	$\mathbf{R} : \mathbf{S}$	Double dot product
$\mathbf{K}(X)$	$\mathbf{K}$-structures on X	$[\mathbf{a}^1, \ldots, \mathbf{a}^m]$	Description space
(X, σ)	$\mathbf{K}$-structure	$\lvert\mathbf{R}\rvert$	$= (\mathbf{R} : \mathbf{R})^{1/2}$, norm of $\mathbf{R}$
$\mathbf{K}(\sigma, \tau)$	$\mathbf{K}$-admissible $\mathbf{C}$-morphisms	L^2	Square-integrable functions
$\mathbf{K}$	Category associated with $\mathbf{K}$	$\mathbf{R}$	Category of description spaces
S	Organismic set		
e_i	Element of organismic set	**Vect**	Category of vector spaces
S_i^a	Activities of e_i	$\mathbf{R}(\mathbf{a})$	$= \mathbf{R}$ in $[\mathbf{a}^1, \ldots, \mathbf{a}^m]$
S_i^p	Products of e_i		

I. INTRODUCTION

> "Where shall I begin, please your Majesty?" he asked. "Begin at the beginning," the King said, very gravely, "and go on till you come to the end: then stop."
>
> Lewis Carroll, *Alice in Wonderland*

The single greatest influence in the development of this work is the book "Fundamentals of Measurement and Representation of Natural Systems" by Professor Robert Rosen (1978, Elsevier North-Holland, Inc., New York).

I have always been more interested in "conceptual" biomathematics than in looking at specific models of a biological process. In other words, my interest is oriented more toward the *logic* of mathematical biology. During the course of many stimulating discussions with Professor Rosen, we talked about the theories of measurement, recognition, discrimination, interactions, bifurcations, and classifications, and above all, he taught me the importance of the idea of *alternate descriptions* (Rosen, 1976) of a system. After having learned to think plurally and after having read the book (Rosen, 1978), it became quite natural to me to work on the problems of measurement, interaction, and representation from the standpoint of the theory of categories, which is the logical mathematical tool to be used when one studies a collection of objects with similar structures.

In his book *The Scientific Outlook* (1931) Bertrand Russell described the "scientific process" as composed of three main stages; the first consists in observing the significant facts; the second in arriving at a hypothesis which, if it is true, would account for these facts; the third in deducing from this hypothesis further consequences which can be tested by observation. So according to Russell, the act of *observation* is basic to science. Intuitively, the notion of *observable* is attached to that of a concrete procedure for determining the value assumed by the observable of a system at a specific time. The crucial ingredient of any such procedure is a measuring instrument, which forms the basis both for our knowledge of the physical world and for our formulation of models that organize this knowledge and allow us to predict and control.

Rosen (1978) provides a comprehensive theory of observables and the descriptions arising from them. The theory is then applicable to any situation in which objects of interest are labelled by definite mapping processes, measurement in physics, pattern recognition, discrimination, or classification. All these diverse situations share a common character, namely, the generation of numbers (or other kinds of invariants) that serve to label the processes with which they are associated, such that processes are considered "the same" if and only if they bear the same label. This leads to the idea of observable induced equivalence relations.

It is suggested in Rosen (1978) that a formal treatment of systems with observables using category theory would be a fruitful undertaking. The engagement in this problem marked the beginning of my dissertation. The next idea immediately came to mind: Why stop at systems with observables? Since the process of observation ultimately rests on the capacity of a given system to induce a dynamics (i.e., a change of state) in a measuring instrument (alias meter, recognizer, discriminator, classifier, etc.), it seems natural to consider systems with dynamics as well. My definition of *dynamics* is more general than that in Rosen. Roughly, the trajectories of my dynamics are not defined on the whole real line as in Rosen, but are only required to be defined on a (possibly bounded) neighbourhood (a, b) of zero, in which case the "observed value" of the dynamical interaction can simply be defined as the value at b, whereas in the former case one needs asymptotic values.

There is a reciprocity that exists in general dynamical interactions between systems. The process of measurement can be considered as a reciprocal induction of dynamics in both the system being measured and the system that measures. Then the basic problem in the analysis becomes to determine the observables through which a particular given dynamics is taking place, to specify the subsystems to which these observables belong, and to identify the manner in which each of these subsystems is causing the others to change states.

The concepts of linkage of observables, stability, and bifurcation, and their connections with dynamics are introduced in Rosen (1978). In my study I treated them in the context of the category of natural systems, an amalgamation of states, observables, and dynamics. Thus my formal treatment of natural systems was initiated by, but very different from, Rosen's book. Further, the categories I constructed have many curious links with diverse branches of mathematics from topology to Galois theory and provide a natural setting to discuss the modelling relation.

During the course of this work, some questions suggest themselves naturally: the problems of growth, differentiation, development, and aging arose when I considered the evolution of natural systems through time. These problems converged to the concept of the organism, which led to the organismic sets of Nicolas Rashevsky (1972) and the living systems of James Miller (1978). At the time of writing this thesis, I was also collaborating with Professor I. W. Richardson on a paper (Richardson, Louie, and Swaminathan, 1982) in which we developed a phenomenological calculus underlying his theory of the description space. I was pleasantly surprised to find a connection between the ideas of categorical system theory in my thesis and this phenomenological calculus. It is quite remarkable that the apparently completely different methods used by my two teachers, Professors Rosen and Richardson, to analyze complex natural systems are actually very close in essence through this connection woven into my thesis.

From whatever side we approach our principle, we reach the same conclusion.

II. PRELIMINARIES: CATEGORY THEORY

> There is therefore only a single categorical imperative and it is this: Act only on that maxim through which you can at the same time will that it should become a universal law.
>
> Immanuel Kant

1. General Discussion

Many problems in mathematics are not primarily concerned with a single object such as a function, a group, or a measure, but deal instead with large classes of such objects. The classes consist of sets with a given structure and of the mappings preserving this structure.

Thus we may be dealing with groups; the mappings preserving the group structure are the (group-) homomorphisms. When we turn to vector spaces, the appropriate mappings are the linear transformations. And when we study topological spaces, continuous functions arise naturally.

A useful discussion of this situation can be given within a general framework that assumes only very little about the mappings. All we need is that they are closed under composition and include the identity mapping. But we need not even assume them to be mappings. The definition takes the following form.

2. *Category*

DEFINITION. A *category* **C** consists of

(1) a collection of *objects*.

(2) For each pair of **C**-objects A, B, a set $\mathbf{C}(A,B)$, the *hom-set* of *morphisms* from A to B. If $f \in \mathbf{C}(A,B)$, we also write $f: A \to B$.

(3) For any three objects A, B, C a mapping

$$\mathbf{C}(A,B) \times \mathbf{C}(B,C) \to \mathbf{C}(A,C)$$

taking $f: A \to B$ and $g: B \to C$ to its *composite* $g \circ f: A \to C$.

These satisfy the following three axioms:

(i) $\mathbf{C}(A,B) \cap \mathbf{C}(C,D) = \varnothing$ unless $A = C$ and $B = D$. (Thus each morphism $f: A \to B$ determines uniquely its *domain* $A = \operatorname{dom} f$ and its *codomain* B. So the objects in a category **C** are really redundant and one can simply consider **C** as a collection of morphisms, or *arrows*.)

(ii) *Associativity:* If $f: A \to B, g: B \to C, h: C \to D$, (so that both $h \circ (g \circ f)$ and $(h \circ g) \circ f$ are defined), then $h \circ (g \circ f) = (h \circ g) \circ f$.

(iii) *Identity:* For each object A, there exists $1_A: A \to A$ such that for any $f: A \to B$, $g: C \to A$, one has $f \circ 1_A = f$, $1_A \circ g = g$. (It is clear that 1_A is unique.) 1_A is called the *identity morphism* on A.

3. *Isomorphisms*

Speaking informally, two mathematical systems of the same nature are said to be *isomorphic* if there is a one-to-one mapping of one onto the other that preserves all relevant properties, or a "structure-preserving bijection." Such a mapping is an *isomorphism* and it usually coincides with the intuitively most natural concept of structural preservation. Categorically, we have the

DEFINITION. A morphism $f: A \to B$ is an *isomorphism* if there exists an *inverse morphism* $g: B \to A$ such that $g \circ f = 1_A$ and $f \circ g = 1_B$. It is clear that if such inverse morphism exists it is unique and may be denoted by f^{-1}.

An isomorphism with domain and codomain both A is an *automorphism* on A. If there exists an isomorphism from A to B then A and B are *isomorphic*, denoted by $A \cong B$. Isomorphic objects are abstractly the same and most constructions of category theory are only unique "up to isomorphism."

4. Examples of Categories

(1) The category **Ens** of sets and functions (i.e., the class of **Ens**-objects is the class of all sets). (In this study we shall take the naive viewpoint of set theory and assume the existence of a suitable universe of all sets—all *small* sets.) And for sets A and B, $\mathbf{Ens}(A, B) = B^A$ is the set of all functions from A to B. Note that $f: A \to B$ and $f: A \to B'$, where $f(A) \subset B$, $f(A) \subset B'$, and $B \neq B'$, are considered as different **Ens**-morphisms even though as mappings they are the same. It is understood that similar remarks apply to the other examples that follow. Also, **Ens**-isomorphisms are just bijections.

(2) The category **Gp** of groups and homomorphisms.

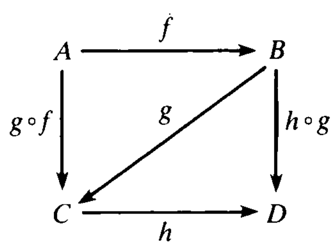
(3) The category **Top** of topological spaces and continuous functions. **Top**-isomorphisms are homeomorphisms.

(4) The category of nonempty sets and functions.

(5) Let M be a monoid. We can regard M as a category with a single object, whose morphisms are the elements of M. Note the identity in the monoid is the identity morphism and composition of morphisms is the monoid operation. If all morphisms are isomorphisms, the monoid is a group, and conversely. (Indeed, if A is any object in a category $\mathbf{C}$, the hom-set $\mathbf{C}(A, A)$ can be considered as a monoid.)

(6) A partially ordered set $(S, \leq)$ may be considered as a category whose objects are the elements of S, and $S(a, b)$ for $a, b \in S$ has a single element or is empty according to whether $a \leq b$ or not.

From these examples we see that morphisms are generally, but not always, mappings. We shall, most of the time, illustrate the composition of morphisms by diagrams as for mappings. For example, the associativity condition of II.2(ii) is equivalent to saying that in



the square commutes.

5. Subcategory

Given categories **A** and **B**, we say that **A** is a *subcategory* of **B** if each **A**-object is a **B**-object, each **A**-morphism is a **B**-morphism, and composition of morphisms is the same in **A** and **B**. Thus for any two **A**-objects X and Y, we have

$$\mathbf{A}(X, Y) \subset \mathbf{B}(X, Y).$$

If equality holds for all X and Y, **A** is a *full subcategory* of **B**.

The category of sets and injections is a subcategory of **Ens** that is not full, while the category of nonempty sets and functions is a full subcategory of **Ens**. The category **Gp** is a full subcategory of the category **Mon** of monoids and monoid-homomorphisms—every monoid-homomorphism between groups is actually a group-homomorphism; and the category **Ab** of abelian groups is a full subcategory of **Gp**.

Note that for a general subcategory we must specify both objects and morphisms and for a full subcategory we need only specify the objects.

6. *Functors*

The idea of a "structure-preserving mapping" can be extended to categories themselves. Let **A** and **B** be categories; then a (*covariant*) *functor* from **A** to **B** is a function $F: \mathbf{A} \to \mathbf{B}$ that assigns to each **A**-object X a **B**-object FX and to each **A**-morphism $f: X \to Y$ a **B**-morphism $Ff: FX \to FY$ such that

(i) If $g \circ f$ is defined in **A**, then $Fg \circ Ff$ is defined in **B** and $F(g \circ f) = Fg \circ Ff$.

(ii) For each **A**-object X, $F1_X = 1_{FX}$.

Besides the covariant functors there is another kind of functor, which reverses the composition. A *contravariant functor* F from **A** to **B** assigns to each **A**-object X a **B**-object FX and to each **A**-morphism $f: X \to Y$ a **B**-morphism $Ff: FY \to FX$ such that (ii) above holds and

(i*) If $g \circ f$ is defined in **A**, then $Ff \circ Fg$ is defined in **B** and $F(g \circ f) = Ff \circ Fg$.

7. *Dual Category*

Associated with each category **C** there is another category called its *dual* or *opposite*, denoted by $\mathbf{C}^{op}$, formed by "reversing all the arrows." Explicitly, $\mathbf{C}^{op}$ has the same objects as **C**, but to each **C**-morphism $f: X \to Y$ there corresponds a $\mathbf{C}^{op}$-morphism $f^{op}: Y \to X$, so that $f^{op} \circ g^{op}$ is defined whenever $g \circ f$ is defined, and that $(g \circ f)^{op} = f^{op} \circ g^{op}$.

Note that a contravariant functor from **A** to **B** may be described as a covariant functor from $\mathbf{A}^{op}$ to **B**, or as a covariant functor from **A** to $\mathbf{B}^{op}$. Also, $(\mathbf{A}^{op})^{op} = \mathbf{A}$.

8. Examples of Functors

(1) There is a covariant functor from **Gp** (or **Top**, or any category of "sets with structure" in general) to **Ens** which assigns to each group (or topological space) its underlying set, and regards each homomorphism (or continuous function) as a mapping of sets. This functor is called the *forgetful functor.*

(2) If **A** is a subcategory of **B**, then the inclusion of **A** in **B** is a functor, the *inclusion functor.*

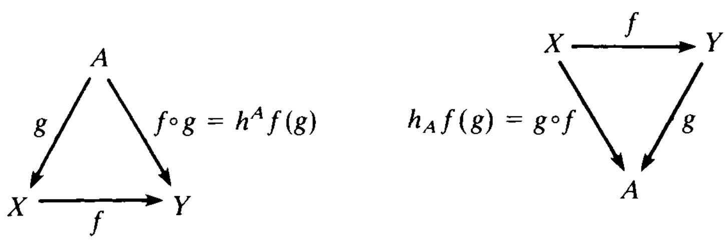
(3) Let M and N be monoids, regarded as categories [see Example II.4(5)]. A covariant functor from M to N is simply a monoid-homomorphism.

(4) For any category **C** and **C**-object A, there is a covariant functor h^A [and a contravariant functor h_A] from **C** to **Ens** that assigns to a **C**-object X the set $h^A X = \mathbf{C}(A, X)$ [and $h_A X = \mathbf{C}(X, A)$] and to a **C**-morphism $f: X \to Y$ the function $h^A f: \mathbf{C}(A, X) \to \mathbf{C}(A, Y)$ [and $h_A f: \mathbf{C}(Y, A) \to \mathbf{C}(X, A)$] defined by

$$h^A f(g) = f \circ g \qquad \text{for} \quad g: A \to X$$

$$[\text{and } h_A f(g) = g \circ f \qquad \text{for} \quad g: Y \to A],$$

i.e., via the diagrams



h^A and h_A are known as the *covariant hom-functor* and the *contravariant hom-functor*, respectively.

(5) The *homology functors* H_n from **Top** to **Ab** that take a topological space X to its singular homology groups $H_n(X)$, and a continuous function $f: X \to Y$ to the homomorphisms $f_*: H_n(X) \to H_n(Y)$ are covariant functors. Similarly, the *cohomology functors* H^n: **Top** $\to$ **Ab** are contravariant functors. In fact, it is in the study of algebraic topology that the ideas of category theory originated (see Eilenberg and Mac Lane, 1945). It is interesting to note that category theory arises from the area of mathematics that gives an example of the power of the modelling relation. (In algebraic topology the topological spaces can be considered as being "modeled" by various algebraic objects. For a discussion of the modelling relation, see Section VI.)

THEOREM. Let F: **A** $\to$ **B** be a (covariant or contravariant) functor. Then F maps **A**-isomorphisms to **B**-isomorphisms. □

Applied to algebraic topology, for example, this theorem says that if $H_n(X)$ and $H_n(Y)$ are not isomorphic groups, then the topological spaces X and Y are not homeomorphic. This is why algebraic topology is used as a tool in dealing with the "homeomorphism problem."

9. *The Category* **Cat**

The idea of category can be applied to categories and functors themselves. Functors can be composed—given functors $F: \mathbf{A} \to \mathbf{B}$ and $G: \mathbf{B} \to \mathbf{C}$ the maps $X \mapsto G(FX)$ and $f \mapsto G(Ff)$ on **A**-objects X and **A**-morphisms f define a functor $G \circ F: \mathbf{A} \to \mathbf{C}$. This composition is clearly associative. For each category **A** there is an *identity functor* $I_{\mathbf{A}}: \mathbf{A} \to \mathbf{A}$. So we may consider the category **Cat** having as objects all *small* categories and as morphisms all functors between them. The reason we have to consider small categories and not any category in general is that we want the **Cat**-objects to be in the universe of set theory (see Example II.4(1)) and hence we want the collection of objects of a **Cat**-object to be a *small set.*

An *isomorphism* $F: \mathbf{A} \to \mathbf{B}$ of categories is a functor that is a bijection both on objects and on morphisms. This is clearly equivalent to the existence of an "inverse functor" $F^{-1}: \mathbf{B} \to \mathbf{A}$.

10. *Full and Faithful Functors*

A functor $F: \mathbf{A} \to \mathbf{B}$ is *full* if to each pair X, Y of **A**-objects and to every **B**-morphism $g: FX \to FY$ there is an **A**-morphism $f: X \to Y$ such that $g = Ff$.

A functor $F: \mathbf{A} \to \mathbf{B}$ is *faithful* if to each pair X, Y of **A**-objects and to every pair $f_1, f_2: X \to Y$ of **A**-morphisms the equality $Ff_1 = Ff_2: FX \to FY$ implies $f_1 = f_2$.

For example, if $F: M \to N$ is a surjective but not injective monoid homomorphism, then when M and N are regarded as categories [Example II.4(5)], F can be considered as a functor and it is full but not faithful. On the other hand, the forgetful functor from **Gp** to **Ens** is faithful but not full.

These two properties may be visualized in terms of hom-sets. For each pair of **A**-objects X and Y, the functor $F: \mathbf{A} \to \mathbf{B}$ assigns to each $f \in \mathbf{A}(X, Y)$ a morphism $Ff \in \mathbf{B}(FX, FY)$, and so defines a function

$$F_{X,Y}: \mathbf{A}(X, Y) \to \mathbf{B}(FX, FY)$$

with $F_{X,Y}(f) = Ff$. Then F is full when every such function is surjective and faithful when every such function is injective. If F is both full and faithful, then every such function is a bijection, but this does not mean that F is a **Cat** isomorphism, for there may be **B**-objects that are not in the image of F.

If **A** is a subcategory of **B**, then the inclusion functor is faithful. It is full if and only if **A** is a full subcategory of **B**.

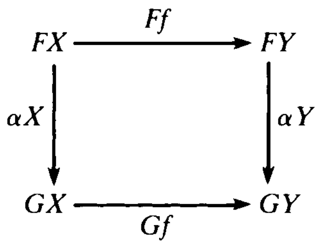
As another example, consider Rosen's (1958) representation theorem of biological systems: "Given any system M and a resolution of M into components, it is possible to find an abstract block diagram which represents M and which consists of a collection of suitable objects and mappings from the category of all sets," i.e., there is a faithful functor from the "category of all systems" to **Ens**.

11. Natural Transformation

DEFINITION. Suppose $F, G: \mathbf{A} \to \mathbf{B}$ are two functors between the same two categories. A *natural transformation* α from F to G is defined by

(1) for each **A**-object X there is a **B**-morphism $\alpha X: FX \to GX$,

(2) for each **A**-morphism $f: X \to Y$ the following square of **B**-morphisms commutes:



We use the notation $\alpha: F \to G$ when α is a natural transformation from F to G.

12. Functor Category

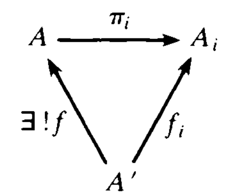
If **A** and **B** are categories, we can define the *functor category* $\mathbf{B}^{\mathbf{A}}$ to have as objects all (covariant) functors from **A** to **B**, to have as morphisms natural transformations, and to have as composition and identities the "pointwise" ones. It is easy to check that $\mathbf{B}^{\mathbf{A}}$ is indeed a category.

Isomorphisms in $\mathbf{B}^{\mathbf{A}}$ are called *natural equivalences*. The natural transformation $\alpha: F \to G$ is a natural equivalence if and only if each αX (for **A**-objects X) is an isomorphism (from FX to GX) in **B**.

13. Products

In set theory, the product of a family $\{A_i : i \in I\}$ of sets is the set $A = \Pi A_i$ of all I-tuples $(a_i : i \in I)$ with each $a_i \in A_i$. The function $\pi_j: A \to A_j$ that sends (a_i) to a_j is called the *jth projection* (or projection onto the jth coordinate). If A' is a set and if there are functions $f_j: A' \to A_j$, then there exists a unique function $f: A' \to A$ such that $\pi_j \circ f = f_j$ for all $j \in I$; namely, $f(a') = (f_i(a') : i \in I)$.

With the preceding motivation, given a family of objects $\{A_i : i \in I\}$ in a category **C**, a *product* of this family in **C** is an object A of **C**, often denoted by ΠA_i, with an I-tuple of **C**-morphisms (*projections*) $\pi_i : A \to A_i$, possessing the *universal property* that whenever A' is a **C**-object similarly equipped with $f_i : A' \to A_i$, there exists a unique **C**-morphism $f: A' \to A$ such that for all $i \in I$, $\pi_i \circ f = f_i$:



A category **C** *has products* if ΠA_i exists for every family $\{A_i : i \in I\}$.

LEMMA. Any two products of $\{A_i : i \in I\}$ are isomorphic. □

This lemma holds for all universal properties, i.e., any constructions in a category via universal properties are unique up to isomorphism.

DEFINITION. The product of the empty family in **C** is the *final object* of **C**. X is the final object of **C** if and only if for every **C**-object A there is a unique **C**-morphism from A to X.

EXAMPLES. **Ens** has products, the usual "cartesian product sets." The final object in **Ens** is a singleton set. **Top** has products, the cartesian product sets with the product topologies. The final object is a singleton set with the only topology. **Gp** has products, the direct products of groups. The final object is the trivial group.

14. Duality

For each concept in a category **C**, there is a *co-concept* from its dual category $\mathbf{C}^{op}$. For example, if $f: A \to B$ is a **C**-morphism, then A is the *domain* of f in **C**; now $f^{op}: B \to A$ in $\mathbf{C}^{op}$ is such that B is the domain of f^{op}, hence B is the *codomain* of f in **C**. This is consistent with Definition II.2(i).

If $\Sigma(\mathbf{C})$ is a statement about an arbitrary category **C**, let Σ^{op} be the statement defined by $\Sigma^{op}(\mathbf{C}) = \Sigma(\mathbf{C}^{op})$. For example, consider the statement in Lemma II.13:

$\Sigma(\mathbf{C})$ = Products in a category **C** are unique up to isomorphism.

Then

$\Sigma(\mathbf{C}^{op})$ = Products in a category $\mathbf{C}^{op}$ are unique up to isomorphism.

Hence we have the Lemma II.13^{op}:

$\Sigma^{op}(\mathbf{C}) = \Sigma(\mathbf{C}^{op})$ = *Coproducts* in a category **C** are unique up to isomorphism.

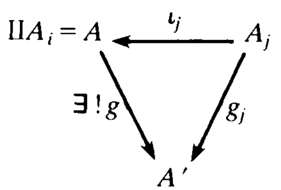
(Note that co-isomorphisms are isomorphisms.)

The Principle of Categorical Duality is: Σ^{op} is *universally* true if Σ is. (Note that "universally true" means that the statement is a consequence of the category axioms.) Duality cuts the work in half.

15. Coproducts

A **C**-object A with **C**-morphisms $\iota_j: A_j \to A$ is the *coproduct* of the family $\{A_i : i \in I\}$ in **C** if, of course, $(A, \iota_j^{op}: A \to A_j)$ is the product of $\{A_i : i \in I\}$ in $\mathbf{C}^{op}$. Explicitly, (A, ι_j) is the coproduct of $\{A_i\}$ in **C** if for $(A', g_j: A_j \to A')$ there exists a unique **C**-morphism $g: A \to A'$ such that for all $j \in I$, $g \circ \iota_j = g_j$. The morphisms ι_j are called *injections* and the coproduct of $\{A_i\}$ is denoted by $\coprod A_i$.

Diagrammatically, we have:



A category **C** *has coproducts* if $\coprod A_i$ exists for every family $\{A_i : i \in I\}$.

The coproduct of the empty family is the *initial object* of **C**. X is the initial object of **C** if and only if for every **C**-object A there is a unique **C**-morphism from X to A.

EXAMPLES. **Ens** has coproducts, the disjoint union $\coprod A_i = \Pi\{i\} \times A_i = \{(i, a) : i \in I, a \in A_i\}$ and the jth injection ι_j sending $a \in A_j$ to (j, a). The unique initial object of **Ens** is the empty set $\varnothing$. (By the Bourbaki convention, $X^{\varnothing} = \{\varnothing\}$ for any set X; i.e., there is a unique function from the empty set to any set, namely, the inclusion map of the empty set into X—the "empty function.")

Top has coproducts, the disjoint union equipped with the direct sum topology—a set G in $\coprod A_i$ is open if and only if $G \cap A_i$ is open for each i. The initial object is the empty space.

The coproduct in categories of sets with structure may have its underlying set different from the disjoint union. For example, in **Ab**, the coproduct $\coprod A_i$ is the subgroup of ΠA_i consisting of all I-tuples (a_i) such that $a_i = e_i$ = the identity of A_i for all but finitely many i. $\coprod A_i$ is usually referred to as the direct sum $\bigoplus A_i$. And for the universal property, $g: \amalg A_i \to A'$ is defined by $g((\bar{a}_i : i \in I)) = \Sigma g_i(a_i)$. Note that $a_i = e_i$ hence $g_i(a_i) = e' =$ the identity of A' for all but finitely many i, so there is no "convergence" problem in the sum. The initial object of **Ab** is the trivial group that is also the final object. In a category **C**, an object that is both initial and final is called the *zero object*.

This ends the preliminaries on the fundamentals of category theory, which is put here for the notation and for completeness. Many other constructions are possible in a category, and they will be introduced and defined when encountered. The reader is referred to any standard text on category theory (e.g., Mac Lane, 1971) for more detailed discussions.

III. THE CATEGORY OF FORMAL SYSTEMS

> To banish reality is to sink deeper into the real; allegiance to the void implies denial of its voidness.
>
> Seng-Ts'an

A. Introduction

1. General Discussion

Throughout this study we will be dealing with three basic undefined terms: *system*, *state*, and *observable*. Intuitively, a *system* is some part of the real world that is our object of study; a *state* is a specification of what our system is like at a particular time; and an *observable* of the system is some characteristic of the system that can, at least in principle, be measured. In other words, an observable of a system is a quantity that can induce dynamics in some appropriate meter.

These three basic concepts are interrelated via two fundamental propositions that we shall take as axioms in all of what follows. These propositions are as follows:

PROPOSITION 1. The only meaningful physical events that occur in the world are those represented by the evaluation of observables on states.

PROPOSITION 2. Every observable can be regarded as a mapping from states to real numbers.

Scientific activity usually starts with the collection of observable phenomena within a given field. The significance of Proposition 1 lies in the word *meaningful*—because our information on a system is derived from what we can observe; hence what we cannot observe (in the generalized sense of creating percepts in our brains) will not be meaningful to us. This points also to the fact that as our means of observation increase, the more "alternate descriptions" we have on a system, we will have more insights in how and why things work.

To see why Proposition 2 makes sense, let us first quote Sir Arthur Eddington from his *Fundamental Theory* (1949): "The whole subject matter of exact science consists of pointer readings and similar indications; whatever quantity we say we are 'observing', the actual procedure nearly always ends in reading the position of some kind of indicator on a graduated scale or its equivalent." In science the most common form of questioning Nature is through experiments, and the answer often comes in *numerical* readings.

It may well be that from the point of view of scientific enquiry, the only way we can handle any entity is through the numbers associated with its measurement, which in turn is defined through the measuring instrument, the meter. This, in particular, is the view of Einstein's on *time*. Such a definition of time—as something dependent of a measuring instrument—is called an *operational definition*. Borrowing the terminology, our two propositions are then saying that all our observables are operationally defined, and all our meters are real-valued.

It must be realized, however, that the operational definition of scientific entities will enable us to investigate only certain aspects of science. Other questions are in principle unanswerable (i.e., they will not yield a measurement that gives an answer to the question). Such questions are then meaningless in the context of scientific investigation. But the study of those aspects amenable to science based on operational definitions is enough to keep us busy forever. After all, the study of a model of the *whole* of Nature belongs to metaphysics, not science.

2. *Equivalence Relations*

Let us consider the prototype situation, in which we have a set S of states and a real-valued function $f: S \to \mathbb{R}$ that represents an observable. f induces an equivalence relation R_f on S defined by

$$s_1 R_f s_2 \qquad \text{iff} \quad f(s_1) = f(s_2).$$

Clearly, the quotient set S/R_f is in one-to-one correspondence with the spectrum $f(S)$. In general the observable f conveys limited information about its domain S, because by definition it cannot distinguish between states lying in the same equivalence class, and the set of states of our system would appear to be S/R_f. This is why "alternate descriptions" of a system are important: The more observables we have, the more information we have on S.

On the other hand, the equivalence relations on S induce an equivalence $\sim$ on $\mathbb{R}^S$, the set of all real-valued functions on S, as follows. For $f, g \in \mathbb{R}^S$, define $f \sim g$ if and only if $R_f = R_g$, i.e., if and only if $f(s_1) = f(s_2)$ is equivalent to $g(s_1) = g(s_2)$. Two equivalent observables "convey the same information" about the elements of S; therefore we cannot distinguish between elements of S by employing equivalent observables. Note, however, that $f \sim g$ only means that $S/R_f = S/R_g$, i.e., there is a one-to-one correspondence between $f(S)$ and $g(S)$, and there need be no relation whatsoever between the values of $f(s)$ and $g(s)$ for $s \in S$. In particular, $|f(s_1) - f(s_2)|$ small does not imply $|g(s_1) - g(s_2)|$ small. Thus if we are considering the metric aspects of observables (which we shall do in later sections), we cannot pass to the equivalence classes in $\mathbb{R}^S/\sim$. But when we are only interested in the *reduced states* in S/R_f and not the specific values of $f(S)$, it is more convenient to consider observables as elements of $\mathbb{R}^S/\sim$. One reason for this is that $\mathbb{R}^S/\sim$ is a partially ordered set under the relation *refinement*. [Define f *refines* g, $f \leq g$, if $f(s_1) = f(s_2)$ implies $g(s_1) = g(s_2)$, i.e., if $R_f \subset R_g$, for $f, g \in \mathbb{R}^S$. Then it is clear that $\leq$ is reflexive and transitive on $\mathbb{R}^S$ and hence is a preorder. But $f \leq g$ and $g \leq f$ only implies $R_f = R_g$ (i.e., $f \sim g$) and not $f = g$. So to make $\leq$ antisymmetric, one passes onto $\mathbb{R}^S/\sim$. Note that $f \leq g$ and $g \leq f$ if and only if $f \sim g$.] It is not uncommon in mathematics to consider equivalence classes of functions instead of the functions themselves—the L^p spaces, for example, are equivalence classes of functions with $f \sim g$ if and only if $f = g$ almost everywhere.

B. The Category S

We shall now undertake a formal treatment of systems, considered as abstract mathematical objects. We shall be studying the properties of the category **S** of (formal) systems and comparing them to those of the category **Ens** of sets.

3. Objects

An object of **S**, a formal system, shall consist of a pair (S, F), where S is a set and F is a set of real-valued functions on S. The elements of S are the states and the elements of F are the observables of the formal system. We shall always assume $0 \in F$ (where 0 is the zero function on S sending all states to the number 0) although for brevity we may sometimes omit 0 when we list the elements of F in specific examples. Thus we have F nonempty in order to avoid "empty set pathologies." The observable 0 is simply "identifying the states in S." Note that $(0)_\sim$ is the greatest element in the partially ordered set $(\mathbb{R}^S/\sim, \leq)$.

4. Morphisms

An **S**-morphism $\phi \in \mathbf{S}((S_1, F_1), (S_2, F_2))$ is a pair of functions $\phi \in \mathbf{Ens}(S_1, S_2)$ and $\phi \in \mathbf{Ens}(F_1, F_2)$, such that for all $f \in F_1$ for all $s, s' \in S_1$, $f(s) = f(s')$ implies $(\phi f)(\phi s) = (\phi f)(\phi s')$, i.e., $sR_f s'$ implies $(\phi s)R_{\phi f}(\phi s')$.

Note that this "compatibility" condition is equivalent to saying for all $G \subset F_1$ for all $s, s' \in S_1$, $sR_G s'$ implies $(\phi s)R_{\phi G}(\phi s')$, where $R_G = \cap R_f : f \in G$ (hence $sR_G s'$ if and only if for all $f \in G$ $f(s) = f(s')$) and $\phi G = \{\phi f : f \in G\} \subset F_2$. This means that for all $G \subset F_1$, ϕ can be considered as a mapping from S_1/R_G to $S_2/R_{\phi G}$.

We always define $\phi 0 = 0$. This is compatible because clearly $0s = 0s'$ implies $0(\phi s) = 0(\phi s')$. Note also that for any observable f, the assignment $\phi f = 0$ is acceptable.

5. Identity

Define $1_{(S,F)} \in \mathbf{S}(S, F), (S, F))$ by for all $s \in S$ $s \mapsto s$ and for all $f \in F$ $f \mapsto f$. (Thus for all $G \subset F$, $G \mapsto G$.) Then clearly $1_{(S,F)}$ satisfies the compatibility condition.

6. Composition

Define composition of morphisms in **S** as simultaneously the compositions on the states and on the observables; i.e., if $\phi : (S_1, F_1) \to (S_2, F_2)$ and $\psi : (S_2, F_2) \to (S_3, F_3)$, define $\psi \circ \phi : (S_1, F_1) \to (S_3, F_3)$ by for every $s \in S_1$ $\psi \circ \phi(s) = \psi(\phi(s))$ and for every $f \in F_1$ $\psi \circ \phi(f) = \psi(\phi f)$. Note for $f \in F_1$ and $s, s' \in S_1$, $sR_f s'$ implies $(\phi s)R_{\phi f}(\phi s')$, which in turn implies $\psi(\phi s)R_{\psi(\phi f)}\psi(\phi s')$; so $\psi \circ \phi$ satisfies the compatibility condition.

Clearly, composition so defined is associative, and for $\phi : (S_1, F_1) \to (S_2, F_2)$, $1_{(S_2,F_2)} \circ \phi = \phi = \phi \circ 1_{(S_1,F_1)}$.

7. Isomorphisms

If $\phi : (S_1, F_1) \to (S_2, F_2)$ and $\psi : (S_2, F_2) \to (S_1, F_1)$ are such that $\psi \circ \phi = 1_{(S_1,F_1)}$ and $\phi \circ \psi = 1_{(S_2,F_2)}$, then it is easy to see that $\phi : S_1 \to S_2$ and $\phi : F_1 \to F_2$ must be bijections (**Ens**-isomorphisms) and that for $f \in F_1$ and $s, s' \in S_1$, $f(s) = f(s')$ if and only if $(\phi f)(\phi s) = (\phi f)(\phi s')$, i.e., for every $G \subset F_1$ $S_1/R_G = S_2/R_{\phi G}$.

Thus isomorphic systems are abstractly the same in the sense that there is a "dictionary" (one-to-one correspondence) between the states and between the observables inducing the "same" equivalence relations on the states. In particular, if F and G are two sets of observables on S and there is a bijection $\phi : F \to G$ such that for all $f \in F$ $f \sim \phi f$, then the two systems (S, F) and (S, G)

are isomorphic with the S-isomorphism $1_S: S \to S$, $\phi: F \to G$. Since categorical constructions are only unique up to isomorphism, in the category S all constructions (S, F) are only "unique up to $\sim$-equivalent observables" (i.e., one can always replace F by an $\sim$-equivalent set of observables G in the above sense) even when the set of states S is held fixed. This last comment is particularly important for all constructions in **S** below.

C. Constructions in S

8. Products

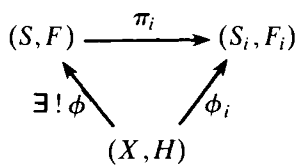
Products in the category **S** do not always exist. For a family $\{(S_i, F_i): i \in I\}$, the product should be $(S, F) = \Pi(S_j, F_j): j \in I$ with an I-tuple of S-morphisms of the form $\pi_i: (S, F) \to (S_i, F_i)$. S is defined as the cartesian product $\Pi\, S_j$ of the sets of states. F is defined as the "cartesian product" $\Pi\, F_j$ of the sets of observables interpreted as follows: for the observable $(f_j: j \in I)$ in F, it is a mapping from S to $\mathbb{R}^I$ defined by

$$(f_j: j \in I)(s_j: j \in I) = (f_j(s_j): j \in I).$$

But S-objects can only have real-valued observables so the mapping $f = (f_j: j \in I): S \to \mathbb{R}^I$ must be *represented* by an equivalent mapping from S to $\mathbb{R}$ (such that S/R_f remains the same). In other words, we need a one-to-one map from $(\mathbb{R}^I)^S$ to $\mathbb{R}^S$ that preserves the equivalence relations on S induced by the observables, or what is equivalent, an injection from $\mathbb{R}^I$ to $\mathbb{R}$. An injection from $\mathbb{R}^I$ to $\mathbb{R}$ only exists when the cardinality of I is less than that of $\mathbb{R}$; thus **S** only has countable (including finite) products. Note, however, that the S-product is independent of the choice of the injection from $\mathbb{R}^I$ to $\mathbb{R}$ because the effect of changing the injection is simply a switch from F to an $\sim$-equivalent set G, and (S, F) and (S, G) are isomorphic (see Section III.7). Different choices of the injection define different S-isomorphism class representatives of the product.

The projections are obviously defined by $\pi_i((s_j: j \in I)) = s_i$ and $\pi_i((f_j: j \in I)) = f_i$. And it is easily checked that the π_i's are indeed S-morphisms.

To see that (S, F) is a product, consider an S-object (X, H) equipped with an I-tuple of S-morphisms $\pi_i: (X, H) \to (S_i, F_i)$. Then



We can define $\phi:(X,H) \to (S,F)$ by for $x \in X$ $\phi(x) = (\phi_j(x): j \in I)$, and for $h \in H$ $\phi(h) = (\phi_j(h): j \in I)$, where again $\phi(h)$ is to be represented by a mapping from S to $\mathbb{R}$ via the injection from $\mathbb{R}^I$ to $\mathbb{R}$. It is clear that ϕ is the unique map that makes the diagram commute. [But ϕ is of course dependent on the choice of the injection from $\mathbb{R}^I$ to $\mathbb{R}$ that determines the product (S,F).] To see that ϕ is an S-morphism, let $h \in H$ and $x, x' \in X$ be such that $h(x) = h(x')$. Then for each $i \in I$ $(\phi_i h)(\phi_i x) = (\phi_i h)(\phi_i x')$ because each ϕ_i is an S-morphism. Whence by definition $(\phi h)(\phi x) = (\phi h)(\phi x')$. Thus $xR_h x'$ implies $(\phi x)R_{\phi h}(\phi x')$. So (S,F) with the π_i's satisfy the universal property in the definition of a categorical product.

The final object in **S** is $(1,\{0\})$ where 1 is the singleton set, the final **Ens**-object. The unique S-morphism from any system to $(1,\{0\})$ is clearly the one that sends all states to 1 and all observables to 0.

9. An Application: Linkage

DEFINITION. Let S be a set of states and $f, g \in \mathbb{R}^S$ be observables. Let p_f: $S \to S/R_f$ and $p_g: S \to S/R_g$ be the natural quotient maps. For $(s)_f \in S/R_f$ consider the set of R_g-classes which intersect $(s)_f$, i.e., the set

$$p_g p_f^{-1}(s)_f = \{(s')_g : f(s') = f(s)\} = \{(s')_g : (s')_g \cap (s)_f \neq \varnothing\}.$$

Then we say

(1) g is *totally linked* to f at $(s)_f$ if the preceding set consists of a single R_g-class (necessarily $(s)_g$; i.e., $f(s) = f(s')$ implies $g(s) = g(s')$);

(2) g is *partially linked* to f at $(s)_f$ if this set consists of more than one R_g-class, but is not all of S/R_g;

(3) g is *unlinked* to f at $(s)_f$ if this set is S/R_g.

Also, we say g is *totally linked* to f if it is totally linked at each $(s)_f$ and g is *unlinked* to f if it is unlinked at each $(s)_f$.

From the preceding definition it is immediate that g is totally linked to f if and only if R_f refines R_g, which is equivalent to the existence of an S-morphism from $(S,\{f\})$ to $(S,\{g\})$ which sends each $s \in S$ to itself and sends f to g, for the latter statement means precisely that $f(s) = f(s')$ implies $g(s) = g(s')$.

For a set of states S equipped with two distinct observables f and g, there is another equivalence relation on S other than R_f and R_g that is of interest—namely, the intersection $R_{fg} \doteq R_f \cap R_g$. The relation R_{fg} is defined by $sR_{fg}s'$ if and only if $f(s) = f(s')$ *and* $g(s) = g(s')$. Note that there may not be an observable of S that generates the equivalence relation R_{fg}, i.e., although mathematically there exists $h \in \mathbb{R}^S$ such that $R_{fg} = R_h$, the set of all possible observables of S, as a representation of a natural system, may not be all of $\mathbb{R}^S$. Thus R_{fg} is generally a formal construction.

There is always an embedding $\phi: S/R_{fg} \to S/R_f \times S/R_g$ that maps $(s)_{fg} \mapsto ((s)_f, (s)_g)$. Via this embedding, a state $s \in S$ is represented by the pair of numbers $(f(s), g(s))$. This embedding ϕ is in general one to one, but it is onto if and only if f and g are totally unlinked (to each other).

This product representation can be constructed neatly as a categorical product. Consider the two systems $(S, \{f, 0\})$ and $(S, \{g, 0\})$. The S-product of these two systems is $(S \times S, F)$, where $F = \{0, (f, 0), (0, g), (f, g)\} \subset \mathbb{R}^{S \times S}$, with the natural projections. Now consider further the system $(S, \{f, g\})$. There exist S-morphisms

$$\phi_1: (S, \{f, g\}) \to (S, \{f, 0\})$$

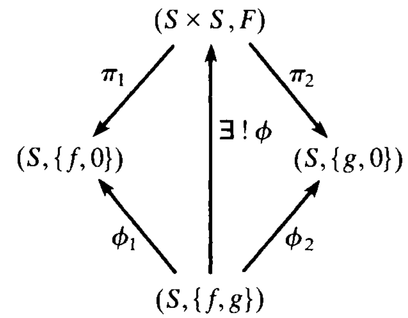
$$\text{and} \quad \phi_2: (S, \{f, g\}) \to (S, \{g, 0\})$$

defined by for every $s \in S$

$$\phi_1(s) = s; \qquad \phi_1 f = f, \quad \phi_1 g = 0$$

$$\text{and} \quad \phi_2(s) = s; \qquad \phi_2 f = 0, \quad \phi_2 g = g.$$

So we have the following diagram:



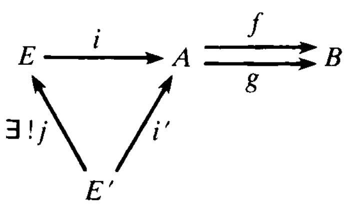
Hence by the universal property of the product, there exists a unique $\phi: (S, \{f, g\}) \to (S \times S, F)$ that makes the diagram commute. Namely, ϕ is defined by sending $s \in S$ to $\phi(s) = (s, s)$—the diagonal map—and by $\phi f = (f, 0)$, $\phi g = (0, g)$. In particular, ϕ being an S-morphism implies that $\phi: S/R_{\{f,g\}} \to S \times S/R_{\{(f,0),(0,g)\}}$. It is clear that ϕ is a one-to-one mapping (on S) and that $R_{\{f,g\}} = R_{fg}$ hence $S/R_{\{f,g\}} = S/R_{fg}$. Also, $S \times S/R_{\{(f,0),(0,g)\}} \cong S/R_f \times S/R_g$. Thus ϕ is indeed the one-to-one map from S/R_{fg} to $S/R_f \times S/R_g$, and that the degree of onto-ness of ϕ is an indication of the lack of linkage between f and g. The onto-ness of a morphism is discussed in Section III.14.

10. Equalizers

Let A and B be sets and $f, g: A \to B$ be two functions. The inclusion map i of the subset $E = \{x \in A : f(x) = g(x)\}$ of A can be characterized up to isomorphism by the following universal property: for every function $i': E' \to A$ such that $f \circ i' = g \circ i'$, there exists a unique function $j: E' \to E$ such that $i \circ j = i'$ (since the image of i' is contained in E, j is defined by $j(x) = i'(x)$ for $x \in E'$).

Generalizing this idea, given morphisms f, $g: A \to B$ in a category **C**, an *equalizer* of (f, g) is an object E with a morphism $i: E \to A$ satisfying the following universal property:

(1) $f \circ i = g \circ i$;
(2) given i' with $f \circ i' = g \circ i'$, there exists unique j such that $i \circ j = i'$.



(An equivalent way of defining an equalizer is to say that it is a final object in the category of all **C**-morphisms that satisfy $f \circ i = g \circ i$.) As usual for "universal property" definitions, equalizers are unique up to isomorphism and will be denoted by $\mathrm{eq}(f, g)$. A category **C** *has equalizers* if $\mathrm{eq}(f, g)$ exists for every pair $f, g: A \to B$.

For **S**-morphisms ϕ, $\psi: (S_1, F_1) \to (S_2, F_2)$, $\mathrm{eq}(\phi, \psi) = (E, H)$ may not exist. The equalizer would have to be given by $E = \{s \in S_1 : \phi s = \psi s\}$, $H = \{f|_E : f \in F_1,\ \phi f = \psi f\}$ and $\imath: (E, H) \to (S_1, F_1)$ would be the inclusion morphism. But $\imath(f|_E) = f$ may not be uniquely defined because there may be another $g \in F_1$ such that $g|_E = f|_E$ and $\phi g = \psi g$. Thus an **S**-equalizer only exists when the inclusion from H to F_1 is a single-valued function.

Note when $(E, H) = \mathrm{eq}(\phi, \psi)$ does exist, $\imath: (E, H) \to (S_1, F_1)$ has the property that for all $s, s' \in E$ and for all $g \in H$, $g(s) = g(s')$ if and only if $(\imath g)(s) = (\imath g)(s')$, i.e., $E/R_g \cong \imath(E)/R_{\imath g}$. Further, any **S**-morphism $\chi: (X_1, G_1) \to (X_2, G_2)$ that is one-to-one on the states and observables and that has this property (that $X_1/R_g \cong \chi(X_1)/R_{\chi g}$ for all $g \in G_1$) is an equalizer. It is easy to construct a pair of S-morphisms ϕ_1, ϕ_2 with domain (X_2, G_2) such that $(X_1, G_1) = \mathrm{eq}(\phi_1, \phi_2)$. Thus although **S** does not have equalizers for every pair of **S**-morphisms, given an **S**-morphism ϕ with the correct properties one can always find a pair of **S**-morphisms for which ϕ is the equalizer.

Equalizers will be discussed further in the sections on hierarchies of **S**-morphisms.

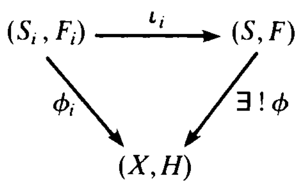
11. Coproducts

S has coproducts only when a cardinality condition is satisfied. The coproduct is $(S, F) = \amalg (S_i, F_i)$ where $S = \amalg S_i$ is the coproduct of the S_i's in **Ens** (i.e., the disjoint union $S = \bigcup \{i\} \times S_i)$ and $F = \{0\} \cup \{(i, f) : i \in I, f \in F_i, f \neq 0\}$ defined as follows. For $f \in F_i$, $f \neq 0$, the observable (i, f) of S is defined by

$$(i, f)(j, s) = \begin{cases} f(s) & \text{if} \quad j = i \\ (j, s) \in \mathbb{R} & \text{if} \quad j \neq i \quad (*), \end{cases}$$

where (*) denotes that we assume there is a one-to-one mapping from $\amalg S_j$: $j \neq i$ to $\mathbb{R}$, i.e., we assume the existence of an observable $=$ on $\amalg S_j : j \neq i$. This assumption is only valid when the cardinality of S is no bigger than that of $\mathbb{R}$. Thus **S**-coproducts only exist when this holds. The natural injections are $\iota_i : (S_i, F_i) \to (S, F)$ with $\iota_i(s) = (i, s)$, $\iota_i f = (i, f)$ for $s \in S_i$ and $f \in F_i$ with $f \neq 0$, and $\iota_i(0) = 0$.

To check the universal property, consider the following:



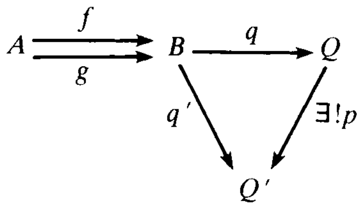
The natural definition for ϕ (given $((X, H), \phi_i)$) is $\phi(i, s) = \phi_i(s)$, $\phi(i, f) = \phi_i f$. Now suppose $(i, f) \in F$, (j, s), $(k, t) \in S$ are such that $(i, f)(j, s) = (i, f)(k, t)$. Then *either* $i = j = k$ and $f(s) = f(t)$ *or* $(j, s) = (k, t)$. If the former, then since ϕ_i is an **S**-morphism, $(\phi_i f)(\phi_i s) = (\phi_i f)(\phi_i t)$, i.e., by definition $(\phi(i, f))(\phi(j, s)) = (\phi(i, f))(\phi(k, t))$. If the latter, then $j = k$ and $s = t$ whence $\phi_j s = \phi_j t$, implying $(\phi_i f)(\phi_j s) = (\phi_i f)(\phi_j t)$, hence $(\phi(i, f))(\phi(j, s)) = (\phi(i, f))(\phi(k, t))$. Thus ϕ is an **S**-morphism.

S, therefore, has an initial object. The initial object in **Ens** is the empty set $\varnothing$, thence the initial object in **S** is $(\varnothing, \{0\})$. For any system (S, F), the unique **S**-morphism from $(\varnothing, \{0\})$ to (S, F) is clearly the empty function on $\varnothing$ with $0 \mapsto 0 \in F$.

12. Coequalizers

$q: B \to Q$ in a category **C** is a coequalizer of $f, g: A \to B$ if, of course, $q^{\text{op}} = \text{eq}(f^{\text{op}}, g^{\text{op}})$ in the dual category $\mathbf{C}^{\text{op}}$.

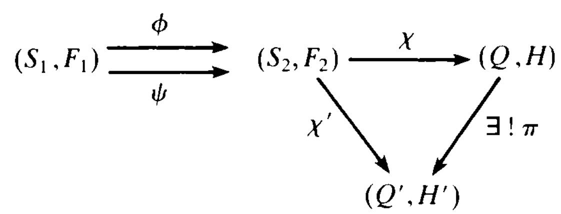
Ens has coequalizers. Given $f, g: A \to B$ let R be the equivalence relation on B generated by $A' = \{(f(a), g(a)) : a \in A\}$, i.e., let R be the intersection of all equivalence relations on B containing A'. Let $Q = B/R$ with canonical projection $q: B \to Q$. Then $q = \mathrm{coeq}(f, g)$. Clearly, $q \circ f = q \circ g$.



And suppose $q': B \to Q'$ and $q' \circ f = q' \circ g$; then $R' = \{(b_1, b_2): q'(b_1) = q'(b_2)\}$ is an equivalence relation on B containing A' and hence contains R; p is thus defined by $p(b)_R = q'(b)$ and so $p \circ q = q'$.

The category **S** also has coequalizers, constructed as follows. Let ϕ, $\psi: (S_1, F_1) \to (S_2, F_2)$. Let $Q = S_2/R$ where R is the intersection of all equivalence relations on S_2 containing $\{(\phi(s), \psi(s)) \in S_2 \times S_2 : s \in S_1\}$ *and* of all R_{F_2}. So in particular for t, $t' \in S_2$, tRt' implies for all $g \in F_2$ $g(t) = g(t')$. Let $\chi: S_2 \to Q$ be the canonical projection $\chi(t) = (t)_R$. This takes care of the map on the states. As for the observables, let R on F_2 be the intersection of all equivalence relations containing $\{(\phi f, \psi f) \in F_2 \times F_2 : f \in F_1\}$, and let $\chi: F_2 \to H = F_2/R$ be, naturally, $\chi g = (g)_R$, where $(g)_R \in \mathbb{R}^Q$ is to be interpreted as follows. Let $(g)_R: S_2 \to \mathbb{R}$ be such that $R_{(g)_R}$ is the equivalence relation on S_2 generated by $\{R_{g'}: g' \in (g)_R\}$, i.e., $R_{(g)_R}$ is the finest equivalence relation on S_2 such that it is refined by each of the $R_{g'}$, $g' \in (g)_R$. Putting it another way, $R_{(g)R}$ is defined to be the supremum of the family $\{R_{g'}: g' \in (g)_R\}$ in the lattice of all equivalence relations on S_2. It is clear, then, that $R_{(g)_R}$ is refined by R on S_2 and hence $(g)_R$ is well defined on $Q = S_2/R$, so we can consider $(g)_R: Q \to \mathbb{R}$. Finally, to check $\chi: (S_2, F_2) \to (Q, H)$ such defined is indeed an S-morphism, let $g \in F_2$ and t, $t' \in S_2$; then $g(t) = g(t')$ implies $(g)_R(t) = (g)_R(t')$ hence $(g)_R(t)_R = (g)_R(t')_R$. So $tR_g t'$ does imply $\chi(t) R_{\chi g} \chi(t')$. And clearly $\chi \circ \phi = \chi \circ \psi$.

Now if $\chi': (S_2, F_2) \to (Q', H')$ is such that $\chi' \circ \phi = \chi' \circ \psi$, then $\{(t, t') \in S_2 \times S_2 : \chi'(t) = \chi'(t')\}$ is an equivalence relation on S_2 containing R. Thus $\pi(t)_R = \chi'(t)$ is well defined on $Q = S_2/R$. Similarly, $\pi(g)_R = \chi' g$ is well defined on H. Clearly, $\chi' = \pi \circ \chi$ and π is unique.



Finally, we have to check that π is an S-morphism. Note that for every $g' \in (g)_R$ (i.e., $g'Rg$), $\chi'g' = \chi'g$ because $\{(g,g') \in F_2 \times F_2 : \chi'g = \chi'g'\}$ is an equivalence relation on F_2 and since for every $f \in F_1$ $\chi'(\phi f) = \chi'(\psi f)$, this equivalence relation contains all $(\phi f, \psi f)$ and hence contains R. Also since χ' is an S-morphism, for each $g' \in (g)_R$ we have $g'(t) = g'(t')$ implying $(\chi'g')(\chi't) = (\chi'g')(\chi't')$, i.e., $(\chi'g)(\chi't) = (\chi'g)(\chi't')$. Thus $R_{\chi'g}$ is "refined" by each of $R_{g'}$ (on F_2). Since $R_{(g)_R}$ is the supremum of $\{R_{g'} : g' \in (g)_R\}$, we have $R_{(g)_R} \subset R_{\chi'g}$. Therefore $(g)_R(t)_R = (g)_R(t')_R$ in Q implies $(g)_R(t) = (g)_R(t')$ in S_2, which in turn implies that $(\chi'g)(\chi't) = (\chi'g)(\chi't')$ in Q', i.e., $\pi(g)_R[\pi(t)_R] = \pi(g)_R[\pi(t')_R]$ in Q', whence $\pi:(Q,H) \to (Q',H')$ is indeed an S-morphism.

It should be obvious that whereas products and equalizers are easy to define in S, their dual concepts are a lot more complicated. This is observed in many familiar categories. A difficult problem in the study of a specific category is to describe its coproducts and coequalizers ("colimits") explicitly.

D. Hierarchy of S-Morphisms

13. Monomorphisms

Injective is an important property of functions. In particular it is the tool with which one defines *subsets* and the ordering of cardinals. As it turns out, there are many categorical definitions that characterize injective in **Ens**; we will content ourselves with three of them.

Let $f: A \to B$ in a category **C**. The morphism f is *split mono* if there exists a $g: B \to A$ with $g \circ f = 1_A$, and *equalizer* if $f = \mathrm{eq}(g_1, g_2)$ for some pair g_1, g_2: $B \to C$, and *mono* if for all pairs $g_1, g_2: X \to A$ such that $f \circ g_1 = f \circ g_2$, we have $g_1 = g_2$.

It is a general theorem in category theory (the "hierarchy theorem for monomorphisms;" see Manes, 1976) that split monos are equalizers and equalizers are monos. Also, if f and g are mono or split mono, so is $g \circ f$ (when composition is appropriate); if $g \circ f$ is mono or split mono, so is f. (The analogue for equalizers is not always true; it is true, however, when a category has epi-equalizer factorizations. See Section III.16.)

In **Ens**, monos are the same as injective functions, and it is easy to see that all monos are equalizers. If $f: A \to B$ is mono and A is nonempty, then f is split mono (let $g = f^{-1}$ on $f(A)$ and arbitrary elsewhere). Note, however, that the inclusion map of the empty set into a nonempty set is mono, but never split mono (because $B^{\varnothing} = \{\varnothing\}$ for sets B but $\varnothing^B = \varnothing$ for $B \neq \varnothing$). An **Ens**-mono f from A to B will be denoted $f: A \rightarrowtail B$. This notation will sometimes be borrowed for a monomorphism in any category.

What is a mono in the category **S** like? We claim that it is the same as an S-morphism that is injective as set mappings on the set of states and on the set of observables. For suppose $\phi: (S_1, F_1) \to (S_2, F_2)$ is a mono and there are distinct states s and s' in S for which $\phi(s) \neq \phi(s')$, then consider $\psi_1, \psi_2: (S_1, F_1) \to (S_1, F_1)$ with ψ_1 mapping all states in S_1 to s, ψ_2 mapping all states in S_1 to s', and both ψ_1 and ψ_2 acting as identity on F_1. It is easy to check that in this case the S-morphisms ψ_1 and ψ_2 are such that $\phi \circ \psi_1 = \phi \circ \psi_2$ but $\psi_1 \neq \psi_2$, a contradiction. So $\phi: S_1 \to S_2$ must be injective. Also, suppose distinct observables f and f' in F_1 are such that $\phi f = \phi f'$, then consider ψ_1, ψ_2: $(\{s\}, \{f, f'\}) \to (S_1, F_1)$ where $s \in S_1$, ψ_1 is the inclusion, and $\psi_2(s) = s$, $\psi_2 f = f'$, $\psi_2 f' = f$. Again ψ_1, ψ_2 are S-morphisms with $\phi \circ \psi_1 = \phi \circ \psi_2$ but $\psi_1 \neq \psi_2$, a contradiction. So $\phi: F_1 \to F_2$ is also injective. Conversely, it is clear that if an S-morphism $\phi: (S_1, F_1) \to (S_2, F_2)$ is injective on both S_1 and F_1, it is mono.

Now suppose $\phi: (S_1, F_1) \to (S_2, F_2)$ is an equalizer and that S_1 is nonempty. (F_1 is already nonempty because $0 \in F_1$.) Say $\phi = \mathrm{eq}(\psi_1, \psi_2)$ for $\psi_1, \psi_2: (S_2, F_2) \to (X, H)$. Then, as an equalizer, $\phi: (S_1, F_1) \to (S_2, F_2)$ is isomorphic to an inclusion (see Section III.10). So in particular for $f \in F_1$ and $s, s' \in S_1$, $f(s) = f(s')$ if and only if $(\phi f)(\phi s) = (\phi f)(\phi s')$, i.e., $S_1/R_f \cong \phi(S_1)/R_{\phi f}$. Thus ϕ^{-1} is well defined on $\phi(S_1)$ and $\phi(F_1)$ and can be extended to an S-morphism on (S_2, F_2). (We need a nonempty S_1 for the same reason as in **Ens**.) So in **S**, an equalizer with nonempty domain is split mono.

In the examples of Section III.15 we will show that a mono is not necessarily an equalizer, so the hierarchy for monomorphisms in **S** is (for $\phi: (S_1, F_1) \to (S_2, F_2)$ with nonempty S_1)

$$\text{split mono} \Leftrightarrow \text{equalizer} \Rightarrow \text{mono} \Leftrightarrow \text{injection (on both } S_1 \text{ and } F_1).$$

14. Epimorphisms

The dual concepts to split mono, equalizer, and mono are, respectively *split epi*, *coequalizer*, and *epi*. Explicitly, $f: A \to B$ (in a category **C**) is *split epi* if there exists a $g: B \to A$ such that $f \circ g = 1_B$; f is a *coequalizer* if $f = \mathrm{coeq}(g_1, g_2)$ for some pair $g_1, g_2: C \to A$; f is *epi* if for all pairs $g_1, g_2: B \to X$, $g_1 \circ f = g_2 \circ f$ implies $g_1 = g_2$.

Dually, the hierarchy theorem for epimorphisms states that split epis are coequalizers and coequalizers are epis. Also, if f and g are epi or split epi, so is $g \circ f$; if $g \circ f$ is epi or split epi, so is g. (Again, the analogue for coequalizers is not always true, but is true when the category has coequalizer-mono factorizations; see Section III.15.)

In **Ens**, all epis split and all three concepts mean "surjective": since any function $f: A \to B$ composes equally with χ_B and $\chi_{f(A)} \in 2^B$, it is clear that epis are surjective. The axiom of choice says that surjections are split epi. An **Ens**-

epi (and sometimes a C-epimorphism) will be denoted $f: A \twoheadrightarrow B$.

In S, an epi is the same as an S-morphism that is surjective on both the set of states and the set of observables. For suppose $\phi:(S_1, F_1) \to (S_2, F_2)$ is an epi and there is an $s \in S_2 \sim \phi(S_1)$, then $\psi_1, \psi_2:(S_1, F_1) \to (\{0,1\},\{0\})$, where $\psi_1 = \chi_{\phi(S_1)}$ on S_2, $\psi_1 f = 0$ for all $f \in F_2$, $\psi_2 = \chi_{S_2}$ on S_2, $\psi_2 f = 0$ for all $f \in F_2$, provide a pair of S-morphisms such that $\psi_1 \circ \phi = \psi_2 \circ \phi$ but $\psi_1 \neq \psi_2$, a contradiction. So $\phi(S_1) = S_2$. Now suppose there is an $f \in F_2 \sim \phi(F_1)$, then $\psi_1, \psi_2:(S_2, F_2) \to (S_2, F_2)$, where $\psi_1 = 1_{(S_2, F_2)}$, $\psi_2 = 1_{S_2}$ on S_2 and $\psi_2 f = 0$ for all $f \in F_2$, is an example in which $\psi_1 \circ \phi = \psi_2 \circ \phi$ but $\psi_1 \neq \psi_2$, again a contradiction. Thus $\phi(F_1) = F_2$. Conversely, it is clear that an S-morphism $\phi:(S_1, F_1) \to (S_2, F_2)$ that is onto both S_2 and F_2 is epi.

Thus in S, we have

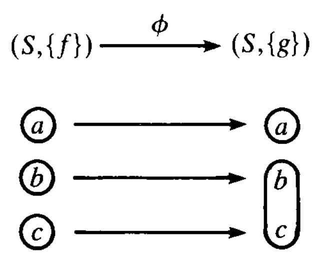
split epi $\Rightarrow$ coequalizer $\Rightarrow$ epi $\Leftrightarrow$ surjection (onto both S_2 and F_2).

In the next section we show that the two preceding one-way implications are indeed irreversible; so the preceding is the hierarchy for epimorphisms in S.

Note that although the two hierarchies in S for the dual concepts of monomorphisms and epimorphisms are not the same, this is *not* a counterexample to the principle of categorical duality (see Section II.14). The principle only states that if Σ is a statement about a category C, then Σ^{op} is *universally* true if Σ is. For a *particular* category, it may very well happen that Σ is true but Σ^{op} is not.

15. Two Examples

Let $S = \{a,b,c\}$, $f: S \to \mathbb{R}$ with $f(a) = 0$, $f(b) = 1$, $f(c) = 2$; i.e., $S/R_f = \{\{a\},\{b\},\{c\}\}$, and $g: S \to \mathbb{R}$ with $g(a) = 0$, $g(b) = g(c) = 1$; i.e., $S/R_g = \{\{a\},\{b,c\}\}$. Let $\phi:(S,\{f\}) \to (S,\{g\})$ be the identity on S, and $\phi f = g$. So diagrammatically we have



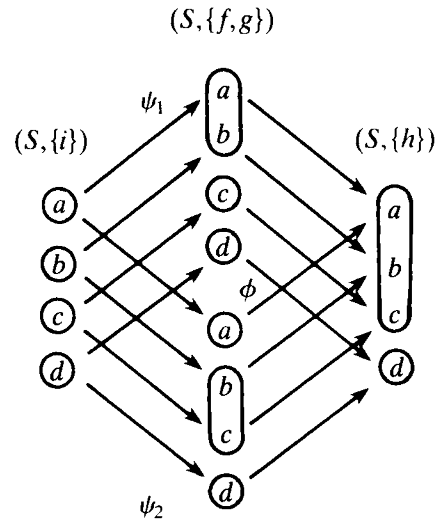
Clearly, ϕ is both mono and epi. Equally obviously, there is only one way to define ϕ^{-1}—namely, ϕ^{-1} is the identity on S, and $\phi^{-1} g = f$. But such ϕ^{-1} would not be acceptable as an S-morphism because b and c are related in $(S,\{g\})$ and not related in $(S,\{f\})$. So this example shows that, in S,

(i) mono and epi $\not\Rightarrow$ isomorphism;
(ii) mono $\not\Rightarrow$ split mono (hence mono $\not\Rightarrow$ equalizer);
(iii) epi $\not\Rightarrow$ split epi.

Further, in this example, since ϕ is mono, any $\psi_1, \psi_2:(X,H) \to (S,\{f\})$ with $\phi \circ \psi_1 = \phi \circ \psi_2$ implies $\psi_1 = \psi_2$. But $\mathrm{coeq}(\psi_1,\psi_1) = 1_{(S,\{f\})} \neq \phi$, so this example also shows that, in **S**,

(iv) epi $\Rightarrow$ coequalizer.

As a second example, consider $S = \{a,b,c,d\}$, $S/R_i = \{\{a\},\{b\},\{c\},\{d\}\}$, $S/R_f = \{\{a,b\},\{c\},\{d\}\}$, $S/R_g = \{\{a\},\{b,c\},\{d\}\}$, $S/R_h = \{\{a,b,c\},\{d\}\}$; and $\psi_1:(S,\{i\}) \to (S,\{f,g\})$ with identity on S, $\psi_1 i = f$, $\psi_2:(S,\{i\}) \to (S,\{f,g\})$ with identity on S, $\psi_2 i = g$, $\phi:(S,\{f,g\}) \to (S,\{h\})$ with identity on S, $\phi f = \phi g = h$. Diagrammatically,



It is easy to check that $\phi = \mathrm{coeq}(\psi_1,\psi_2)$. There can be no $\psi:(S,\{h\}) \to (S,\{f,g\})$ such that $\phi \circ \psi = 1_{(S,\{h\})}$, because we must have either $\psi h = f$ or $\psi h = g$, neither of which is acceptable for ψ to be an **S**-morphism. This shows that ϕ is not split epi. Thus in **S**,

(v) coequalizer $\Rightarrow$ split epi.

E. Images and Subobjects

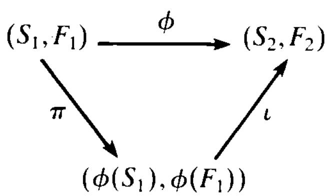
16. Image Factorization

The categorical view of the *image* of a function f is as a factorization $f = i \circ p$, where p is surjective and i is injective. Two of the many possible such views in an arbitrary category are as follows. Given a morphism $f: A \to B$ in **C**, an *epi-equalizer factorization* of f is $f = i \circ p$ with p an epi and i an equalizer. The dual concept is a *coequalizer-mono factorization* $f = i \circ p$ with p a coequalizer and i a mono. It is a general theorem that epi-equalizer [and coequalizer-mono] factorizations are unique up to isomorphism. As a corollary, f is an isomorphism if and only if f is both an equalizer and epi and if and only if f is both a coequalizer and mono. (Note that a morphism that is epi and mono need not be an isomorphism, as we saw in Section III.15 for the case of **S**.)

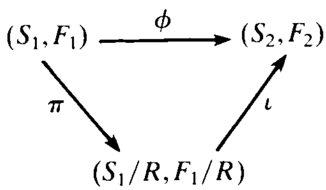
We say **C** *has epi-equalizer* [dually, *has coequalizer-mono*] *factorizations* if every morphism in **C** has an epi-equalizer [respectively, a coequalizer-mono] factorization. The category **Ens** has epi-equalizer and coequalizer-mono factorizations and they both coincide with surjective–injective factorizations.

As a second example, the category **Top** of topological spaces and continuous functions has epi-equalizer and coequalizer-mono factorizations. Given a continuous function $f: A \to B$ with image factorization $f = i \circ p$ at the level of **Ens** we can provide $f(A)$ with the subspace topology induced by i, then (p, i) is an epi-equalizer factorization, or we can provide $f(A)$ with the quotient topology induced by p, in which case (p, i) is a coequalizer-mono factorization. The two image factorizations are clearly not in general homeomorphic.

The category **S**, also, has epi-equalizer and coequalizer-mono factorizations. The diagram



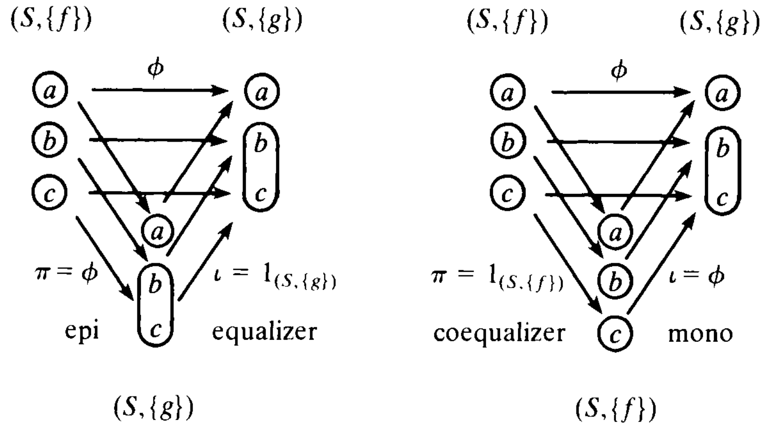
with $\pi = \phi$ on (S_1, F_1) and $\iota =$ inclusion is clearly an epi-equalizer factorization $\phi = \iota \circ \pi$ of ϕ (see Section III.10 that ι may not be well defined). Consider the equivalence relation $R = \{(s, s') \in S_1 \times S_1 : \phi(s) = \phi(s')\}$ on S_1 and $R = \{(f, f') \in F_1 \times F_1 : \phi f = \phi f'\}$ on F_1; the diagram



with $\pi =$ natural projection (where $F_1/R \subset \mathbb{R}^{S_1/R}$ is to be interpreted as in the construction of coequalizers in Section III.12) and ι defined by $\iota(s)_R = \phi(s)$,

$\iota(f)_R = \phi f$ (since sRs' iff $\phi s = \phi s'$, fRf' iff $\phi f = \phi f'$, ι is well defined) is then a coequalizer-mono factorization of ϕ.

Note that although both the epi-equalizer and coequalizer-mono factorizations are unique up to isomorphism, the two factorizations are *not* necessarily isomorphic, as we saw in the case of **Top**. The following example shows this for **S** (same ϕ as in the first example of Section III.15):



It is interesting to compare the image factorizations in **Top** and **S**.

17. Subobjects

In **Ens** (and in most of the categories studied in linear algebra), since a mono splits and all three definitions of monomorphism coincide with injection, there is only one way to define a subobject—namely, we say A is a subset of B (up to isomorphism, of course) if there is an injection $f: A \to B$. In a category **C**, the most common definition of a subobject of an object B is (the isomorphism class of) an object A with a monomorphism $f: A \to B$.

In **S**, as we saw, there are two distinct types of monomorphisms—namely, that of an equalizer (= split mono) and that of a mono (= injective morphism). We shall say that $\phi: (S_1, F_1) \to (S_2, F_2)$ is an **S**-*subsystem* (or simply (S_1, F_1) is a subsystem of (S_2, F_2)) if ϕ is an equalizer, and it is an *S-monosubobject* (or simply (S_1, F_1) is a monosubobject of (S_2, F_2)) if ϕ is mono. So a subsystem is a monosubobject but not vice versa. Note that *subsystem* implies that for each $f \in F_1$ $S_1/R_f \cong \phi(S_1)/R_{\phi f}$ ($\phi: S_1/R_f \to S_2/R_{\phi f}$ is one-to-one), i.e., $f(s) = f(s')$ if and only if $(\phi f)(\phi s) = (\phi f)(\phi s')$; whereas *monosubobject* does not have this "backward implication" ($\phi: S_1/R_f \to S_2/R_{\phi f}$ is not necessarily one to one). A subsystem, therefore, is the appropriate subobject of a system that preserves most of its structures.

18. A Partial Order in **S**

DEFINITION. Let (S_1, F_1).and (S_2, F_2) be **S**-objects. Then $(S_1, F_1) \leq (S_2, F_2)$ if

(i) (S_1, F_1) is a monosubobject of (S_2, F_2), i.e., there is a mono $\phi: (S_1, F_1) \to (S_2, F_2)$;
(ii) S_1 and F_1 are finite sets.

It is clear that $\leq$ is reflexive on *finite* **S**-*objects*, i.e., on (S, F), where S and F are finite sets, and that $\leq$ is transitive. Now suppose $(S_1, F_1) \leq (S_2, F_2)$ (with mono ϕ) and $(S_2, F_2) \leq (S_1, F_1)$ (with mono ψ), then $\phi: S_1 \to S_2$ and $\phi: F_1 \to F_2$ are onto because they are injections between finite sets of the same cardinality. So $\phi: S_1 \to S_2$ and $\phi: F_1 \to F_2$ are **Ens**-isomorphisms. Further, since ϕ and ψ are *order-preserving* on the linkage relations on the finite sets F_1 and F_2, we must have $S_1/R_f \cong S_2/R_{\phi f}$ for every $f \in F_1$. Thus (S_1, F_1) is isomorphic to (S_2, F_2) and so $\leq$ is antisymmetric (up to isomorphism). Therefore $\leq$ is a partial order on (the isomorphism class of) the finite **S**-objects.

What can we deduce from the statement $(S_1, F_1) \leq (S_2, F_2)$? First, we have $S_1 \rightarrowtail S_2$ and in fact for every $s \in S_1$ and every $f \in F_1$ $(s)_f \rightarrowtail (\phi s)_{\phi f}$. So there is a possibility of new states appearing in the whole set and/or in each equivalence class. This reflects *growth* in some respect. The possibility that $F_2 \sim \phi(F_1)$ is nonempty indicates the *emergence* of more observables as the system becomes "more advanced." In particular, there is the possibility that $sR_{F_1}s'$ in S_1 but there is a $g \in F_2 \sim \phi(F_1)$ such that $g(\phi s) \neq g(\phi s')$; so states indistinguishable before could be separated—a model of *differentiation*. On the other hand, it could happen that $f(s) \neq f(s')$ in (S_1, F_1) but $(\phi f)(\phi s) = (\phi f)(\phi s')$ in (S_2, F_2)—a model of *integration* or fusion. Also, since in this case 'distinct' states become indistinguishable, there is an indication of *decay*, or "loss of recognition abilities."

So it seems that with an appropriate totally ordered subset of this partially ordered set of finite systems, a model of the development–senescence process could be constructed. This will be a subject of study in Section VII.

As a final note, condition (ii) in the definition of $\leq$ (that S_1 and F_2 are to be *finite* sets) looks like a very severe mathematical restriction. But in mathematical modelling of natural systems, a finiteness restriction is not unrealistic: All we require is that the sets are finite, and there is no restriction on how small the sets have to be. So the sets could be singletons, have 10^{10} elements, or have 10^{100} elements and still be finite. After all, Sir James Jeans (1945) defined the universe as a gigantic machine whose future is inexorably fixed by its state at any given moment, that it is "a self-solving system of $6N$ simultaneous differential equations, where N is Eddington's number." Sir Arthur Eddington

(1939) asserted (perhaps with more poetry than truth) that $N = 2 \times 136 \times 2^{256}$ ($\sim 10^{79}$) is the total number of particles of matter in the universe. The point is that it is a *finite* number. Thus the set of states of a natural system is certainly finite at one time (this is not to be confused with the set of all *possible* states a system can have), and the set of observables on a system at one time is also clearly finite.

IV. DYNAMICAL SYSTEMS

> The world is a sojourn of all things. Time is a transversal of all ages. The essence of life is change.
>
> Lao Tse

A. Introduction

So far, the discussion has been entirely static. We now include dynamical considerations.

First, we shall restrict our attention to the dynamics of special physical systems called *meters*, through which our observables are defined. Then we shall see that general dynamics, i.e., a change of state in arbitrary systems through interaction of their states, is a corollary of the properties of meters and our fundamental hypothesis (Proposition 1, Section III.1) that all physical events can be represented by the evaluation of observables on states.

1. Meters and Observables

A *meter* M is a set in which a dynamics can be induced through interaction with the states of a system S. By hypothesis, M is in a particular reference state m_0, and M is measuring a specific observable f of S, in the following sense. On interaction of m_0 with a state $s \in S$, a change of state (i.e., a dynamics) is induced in the meter, carrying M (possibly asymptotically) to a state $m(s)$, which is labelled with a real number. The number assigned to $m(s)$ is then defined as the *value* $f(s)$ of the observable f of S. Two states $s, s' \in S$ produce the same response in M if and only if $f(s) = f(s')$. So the meter M can be regarded as interacting with the "reduced" states of S/R_f. Thus every meter defines an observable, and conversely for every observable we assume that there exists a meter in terms of which it could be defined.

2. Meter Dynamics

Since a change of state is a physical event, and since the very essence of the properties of a meter is that a change of state occurs as a result of interaction with other systems, it follows from Proposition 1 that every change of state can be regarded as being specified by the evaluation of observables on states and that every dynamical interaction between systems can be locally represented in terms of the dynamics induced on meters. Thus the reciprocal interaction of two meters sets up a situation similar to that of the "coordinate patches" in differential geometry and is the prototype for general dynamics.

A *dynamics* on a set of states S is usually defined as a one-parameter group (with time being the parameter) of automorphisms (i.e., bijections) on S, $\{T_t \in \mathscr{A}(S) : t \in \mathbb{R}\}$. These automorphisms are in general locally determined through identification of local dynamics with meter dynamics, measuring some observable f of another system S'. So there is a local correspondence $(S', \{f\}) \leftrightarrow (S, \{T_t\})$ between S-objects and "systems with dynamics."

In the next few sections we shall study some of the properties of the collection of "systems with dynamics," again this will be in the setting of category theory.

B. Dynamics and D-Objects

We shall first consider the objects of the category **D** of dynamical systems. It will turn out (Theorem IV.5) that our definition of dynamics is a generalization of the usual definition (as in Section IV.2 and Rosen, 1971).

3. *Dynamics*

DEFINITION. A *dynamics* on a set S is a mapping T from a subset of $S \times \mathbb{R}$ into S satisfying the following conditions:

(i) For every $x \in S$ there exist $a_x, b_x \in \bar{\mathbb{R}}$ (= the extended real numbers $[-\infty, +\infty]$) with $-\infty \leq a_x < 0 < b_x \leq +\infty$ such that $T(x, t)$ is defined if and only if $a_x < t < b_x$. (The notation $a_x[T]$ and $b_x[T]$ will be used when the dependence on T is emphasized.)

(ii) The "initial value property" $T(x, 0) = x$ holds for all $x \in S$.

(iii) The "group property" $T(T(x, t_1), t_2) = T(x, t_1 + t_2)$ holds if both $T(x, t_1)$ and also the left *or* right side of the equation is defined.

LEMMA. The preceding conditions (i), (ii), and (iii) are equivalent to (i), (iii), and

(ii*) T maps *onto* S.

Proof. Clearly, (ii) ⇒ (ii*).

Conversely, assuming (i), (iii), and (ii*), take $x \in S$. Then (ii*) implies that there exists a $y \in S$ and there exists a $t \in (a_y, b_y) \subset \mathbb{R}$ such that $x = T(y, t)$. Now using (iii) and (i), $T(x, 0) = T(T(y, t), 0) = T(y, t + 0) = T(y, t) = x$, proving (ii). □

4. Phase Space

It follows from the lemma that the domain of T is nonempty if and only if S is nonempty. In fact, S is the projection onto the first coordinate of the domain of T. Also, S = the range of T; we shall call S the *phase space* of T.

Each set S is the phase space of at least one dynamics, namely the *trivial dynamics* I_S defined by $I_S(x, t) = x$ for all $x \in S$ and all $t \in \mathbb{R}$.

As an example of a nontrivial dynamics, consider an autonomous differential equation (which, incidentally, started the study of dynamical systems)

$$dx/dt = f(x)$$

on an open subset S of $\mathbb{R}^k$, where $f: S \to \mathbb{R}^k$ is continuous and Lipschitz. Define a dynamics T on S as follows: for $x \in S$, let $y(t)$ be the unique solution of the equation that has $y(0) = x$ and maximal open interval (a_x, b_x) in $\mathbb{R}$ as domain; then set $T(x, t) = y(t)$ for $t \in (a_x, b_x)$. It is easy to check that T: dom $T \to S$ such defined satisfies the definition of a dynamics.

5. Solutions and Translations

Analogously to the preceding example, we can define *solutions* of a dynamics T on S as maps, for each $x \in S$, $y_x: (a_x, b_x) \to S$, $y_x(t) = T(x, t)$. Then it follows that $y_x(0) = x$, $y_{T(x,t_1)}(t_2) = y_x(t_1 + t_2)$ for all $x \in S$ and appropriate $t_1, t_2 \in \mathbb{R}$. The image of each y_x, i.e., the set $y_x(a_x, b_x) = \{y_x(t): t \in (a_x, b_x)\}$ is called a *trajectory* of the dynamics. Note that $x_1 = y_{x_2}(t)$ if and only if $x_2 = y_{x_1}(-t)$, so for each $x \in S$ there is only one trajectory passing through it, namely, $y_x(a_x, b_x)$. This is the *unique trajectory property*, or the *principle of causality*. Clearly, the set of solutions $\{y_x : x \in S\}$ determines T uniquely.

On the other hand, fixing t and varying x, we can define *translations* of a dynamics T as, for each $t \in \mathbb{R}$, T_t: dom $T_t \to S$ where dom $T_t = \{x \in S : T(x, t)$ is defined$\}$, $T_t(x) = T(x, t)$. Then $T_0: S \to S$ with $T_0(x) = x$, $T_{t_2} \circ T_{t_1} = T_{t_1 + t_2}$, and $T_t(x) = y_x(t)$ (for all x and appropriate t). Again, it is clear that the translations $\{T_t : t \in \mathbb{R}\}$ determines T completely.

LEMMA. Each translation T_t: dom $T_t \to S$ is injective.

Proof. Suppose $T_t(x) = T_t(y)$, i.e., $T(x, t) = T(y, t)$. Then $x = T(x, 0) = T(T(x, t), -t) = T(T(y, t), -t) = T(y, 0) = y$. Note the second preceding equality uses the fact that both $T(x, t)$ and $T(x, 0)$ are defined, hence so is $T(T(x, t), -t)$; a similar comment goes to the fourth equality. [Here is an illustration of the importance of the *or* in condition (iii) of Definition IV.3. There is a profound difference if (iii) is altered to hold only when *all* terms concerned are defined.] □

THEOREM. If dom $T = S \times \mathbb{R}$, then each translation $T_t: S \to S$ is an automorphism (of sets, i.e., a bijection).

Proof. T_t is injective by the preceding lemma. If $y \in S$ then $x = T(y, -t) = T_{-t}(y)$ is defined, hence $y = T(y,0) = T(T(y,-t),t) = T(x,t) = T_t(x)$. So T_t is surjective. □

So when dom $T = S \times \mathbb{R}$, $\{T_t : t \in \mathbb{R}\}$ is a one-parameter group of (**Ens**-) automorphisms on S, i.e., a dynamics in the usual definition. So our definition of a dynamics can be considered as a generalization of the usual concept.

6. *Bounds*

DEFINITION. The a_x and b_x appearing in Definition IV.3(i) are called the *bounds of T at x.*

These bounds have the following interesting properties.

LEMMA. For every $x \in S$ and for every $t \in (a_x, b_x)$, $a_{T(x,t)} = a_x - t$ and $b_{T(x,t)} = b_x - t$.

Proof. For $x \in S$ and $t \in (a_x, b_x)$, $T(x,t)$ is defined. Now according to Definition IV.3(iii), $T(T(x,t),t')$ is defined if and only if $T(x, t+t')$ is defined [here again is where the *or* in (iii) is crucial]; then by (i), this means the statement

$$a_{T(x,t)} < t' < b_{T(x,t)}$$

is equivalent to

$$a_x < t + t' < b_x,$$

i.e.,

$$a_x - t < t' < b_x - t.$$

Thus $a_{T(x,t)} = a_x - t$ and $b_{T(x,t)} = b_x - t$. □

COROLLARY 1. If $x \in S$ and either $a_x > -\infty$ or $b_x < +\infty$, then the trajectory $y_x(a_x, b_x) = \{T(x,t): a_x < t < b_x\}$ has the cardinality of the continuum.

Proof. If $a_x > -\infty$ [respectively, $b_x < +\infty$], then all $a_{T(x,t)} = a_x - t$ [respectively, all $b_{T(x,t)} = b_x - t$] are distinct for t varying over (a_x, b_x). So all $T(x,t)$ as t varies over (a_x, b_x) must be distinct. □

COROLLARY 2. If T is a dynamics on a countable set S, then dom $T = S \times \mathbb{R}$. □

This last corollary is particularly interesting: on a countable (including finite) phase space S, the generalized and the usual definitions of *dynamics* coincide.

7. Invariance and Relative Dynamics

DEFINITION. If T is a dynamics on a set S, then $X \subset S$ is *invariant under* T (or simply *T-invariant*) if $T(X, \mathbb{R}) \subset X$ (i.e., for every $x \in X$ and for all $t \in (a_x, b_x)$, $T(x, t) \in X$).

Since $X = T(X, \{0\}) \subset T(X, \mathbb{R})$, the definition is equivalent to $T(X, \mathbb{R}) = X$. Note that S itself, the empty set, and any trajectory of T are all invariant under T.

It is clear that if A is a T-invariant subset of S, then the restriction of T to $(A \times \mathbb{R}) \cap \operatorname{dom} T$ is a dynamics T' on A, and for $x \in A$, $a_x[T'] = a_x[T]$, $b_x[T'] = b_x[T]$.

On the other hand, however, if A is a subset of S and if the restriction T' of a dynamics T on S to some subset of $(A \times \mathbb{R}) \cap \operatorname{dom} T$ is a dynamics on A (we shall call T' a *relative dynamics* on A induced by T on S), A is not necessarily T-invariant. All is required here is that for $x \in A$,

$$a_x[T] \leq a_x[T'] < 0 < b_x[T'] \leq b_x[T]$$

and that T' satisfies Definition IV.3. (In particular, $a_x[T']$ and $b_x[T']$ satisfy Lemma IV.6.) Here $\operatorname{dom} T' \subset (A \times \mathbb{R}) \cap \operatorname{dom} T$ but not necessarily equal.

So we have

$$A \text{ is } T\text{-invariant} \Rightarrow T \text{ induces a relative dynamics on } A,$$

but not conversely.

8. Dynamical System

DEFINITION. A **D**-*object*—a *dynamical system*—is a pair (S, D), where S is a set (of states, the phase space) and D is a set of dynamics on S.

We always assume that the trivial dynamics I_S is in D, so in particular D is always nonempty; but again for brevity we shall sometimes omit listing I_S in specific examples. Note that if a dynamics is considered to be imposed on S through interactions with the states of other systems, then I_S can be considered as being imposed on S through interaction with the (only) state in S/R_0 (where 0 is again the zero observable as in Section III).

C. D-Morphisms

We have studied the objects of **D** in some detail. The next question is, "What are the morphisms?" How shall we define a **D**-morphism so that **D**-isomorphic objects are identical in their dynamical behaviour?

9. **D**-*Morphism*

DEFINITION. $\phi \in \mathbf{D}((S_1, D_1), (S_2, D_2))$ if it is a pair of functions $\phi: S_1 \to S_2$ and $\phi: D_1 \to D_2$ (with $\phi I_{S_1} = I_{S_2}$), and such that for every $T \in D_1$ the diagram

$$\begin{array}{ccc} \text{dom } T & \xrightarrow{\phi \times 1_{\mathbb{R}}} & \text{dom } \phi T \\ {\scriptstyle T}\downarrow & & \downarrow{\scriptstyle \phi T} \\ S_1 & \xrightarrow[\phi]{} & S_2 \end{array}$$

commutes [i.e., for every $x \in S_1$ and for all $t \in (a_x[T], b_x[T])$, $\phi(T(x,t)) = \phi T(\phi x, t)$]. Note this implies that for all $T \in D_1$ and for all $x \in S_1$ $a_{\phi x}[\phi T] \le a_x[T] < 0 < b_x[T] \le b_{\phi x}[\phi T]$.

Also note that the diagram

$$\begin{array}{ccc} S_1 \times \mathbb{R} & \xrightarrow{\phi \times 1_{\mathbb{R}}} & S_2 \times \mathbb{R} \\ {\scriptstyle I_{S_1}}\downarrow & & \downarrow{\scriptstyle \phi I_{S_1} = I_{S_2}} \\ S_1 & \xrightarrow[\phi]{} & S_2 \end{array}$$

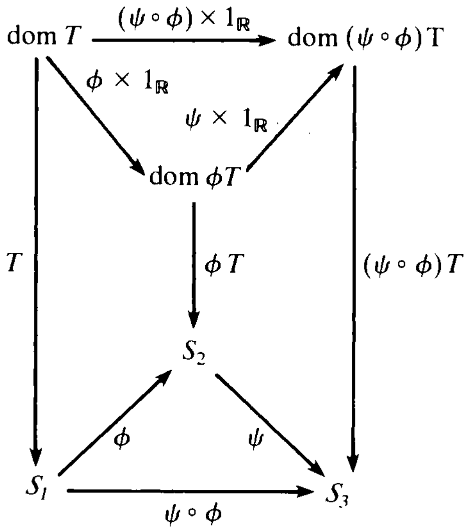
trivially commutes; so the inclusion of I_S into the dynamics does not create any problems.

It follows from the definition of a **D**-morphism that ϕT induces a relative dynamics on $\phi(S_1)$ with relative bounds $a_x[T]$ and $b_x[T]$ at ϕx—in fact, if $T \in D_1$ induces a relative dynamics on $A \subset S_1$, then ϕT induces a relative dynamics on $\phi(A) \subset S_2$. In particular, this last statement holds for $A \subset S_1$ invariant under T (see Section IV.7). Note, however, that A invariant under T does not imply ϕA invariant under ϕT.

Trivially, $1_{(S,D)}: (S, D) \to (S, D)$ sending each $x \in S$ and each $T \in D$ to itself is the identity morphism in $\mathbf{D}((S, D), (S, D))$.

10. *Composition*

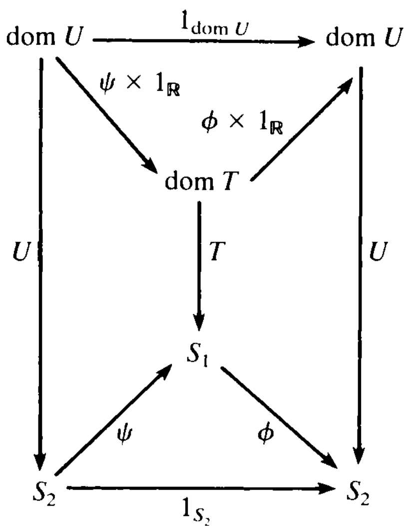
Composition of **D**-morphisms is defined via compositions of functions on the states and on the dynamics. The commutativity condition is represented by the self-explanatory diagram



It is now easy to check that with the preceding definitions, **D** is a category.

11. Isomorphisms

If $\phi:(S_1,D_1) \to (S_2,D_2)$ and $\psi:(S_2,D_2) \to (S_1,D_1)$ are such that $\psi \circ \phi = 1_{(S_1,D_1)}$ and $\phi \circ \psi = 1_{(S_2,D_2)}$, then it follows that $\phi: S_1 \to S_2$ and $\phi: D_1 \to D_2$ are bijections. Further, with $T \in D_1$ and $\phi T = U \in D_2$ (so $\psi U = T$), we have



In particular we see that $U = \phi T = \phi \circ T \circ (\psi \times 1_{\mathbb{R}})$, and for each $x \in S_1$ $a_{\phi x}[\phi T] = a_x[T]$ and $b_{\phi x}[\phi T] = b_x[T]$, hence dom $\phi T \cong$ dom T.

So we see that to say (S_1, D_1) is isomorphic to (S_2, D_2) means that they are the "same" systems down to the fact that corresponding dynamics have isomorphic domains. When dynamical systems are considered as meters, two isomorphic systems are abstractly identical in the sense that they are measuring the same observables. There is a "dictionary" $\phi \leftrightarrow \psi$ that converts one system to the other. This reparametrization can be physically realized as a scale conversion.

D. Constructions in D

12. Products

Let $\{(S_i, D_i): i \in I\}$ be a family in **D**. Then, letting the phase space be $S = \Pi\, S_i$, for each choice of $(T_i: i \in I) \in \Pi\, D_i$, it is natural to define a dynamics T on S—the *product dynamics* of $\{T_i: i \in I\}$—by

$$T((x_i)_{i \in I}, t) = (T_i(x_i, t))_{i \in I}. \tag{1}$$

One can easily check that the canonical projection maps π_j with $x = (x_i)_{i \in I} \mapsto x_j$, $T = (T_i)_{i \in I} \mapsto T_j$ are indeed **D**-morphisms and this construction satisfies the universal property for categorical products.

There is, however, one problem: In Eq. (1), it is required that $T_i(x_i, t)$ be defined for all $i \in I$. It is possible, when I is infinite, that for some choice of the x_i's, we have $a_{(x_i)}[T] = \sup\{a_{x_i}[T_i]: i \in I\} = 0$ and/or $b_{(x_i)}[T] = \inf\{b_{x_i}[T_i]: i \in I\} = 0$. Then in this case T cannot be a dynamics (we need $a < 0 < b$).

So when is $T = (T_i: i \in I)$ a dynamics on S? For $a_{(x_i)}[T] = \sup\{a_{x_i}[T_i]: i \in I\}$ to be negative, it is sufficient that $a_{x_i}[T_i] = -\infty$ for all but finitely many $i \in I$; similarly, for $b_{(x_i)}[T] = \inf\{b_{x_i}[T_i]: i \in I\}$ to be positive, it is sufficient that $b_{x_i}[T_i] = +\infty$ for all but finitely many $i \in I$. Thus when dom $T_i = S_i \times \mathbb{R}$ for all but finitely many i, T is a dynamics on S. In fact, it turns out that this is also a necessary condition:

THEOREM. A necessary and sufficient condition for T defined by Eq. (1) to be a dynamics on $S = \Pi\, S_i$ is that dom $T_i = S_i \times \mathbb{R}$ for all but finitely many $i \in I$.

Proof. The sufficiency is proved above. As for the necessity, suppose $J \subset I$ is a countably infinite subset such that dom $T_j \neq S_j \times \mathbb{R}$ for $j \in J$. Then for each $j \in J$ there exists $y_j \in S_j$ with finite bound(s). Assume, without loss of generality, that $b_{y_j}[T_j] < +\infty$, and that $J = \mathbb{N}$ (i.e., since J is countably infinite, let the corresponding element in $\mathbb{N}$ in the enumeration of J to

$j \in J$ be j). Define $t_j = \max\{0, b_{y_j}[T_j] - 1/j\}$ and $x_j = T_j(y_j, t_j)$. Then by Lemma IV.6

$$b_{x_j}[T_j] = b_{T_j(y_j,t_j)}[T_j] = b_{y_j}[T_j] - t_j \leq 1/j.$$

Now pick $x_i \in S_i$ arbitrarily for $i \notin J$, and let $x = (x_i : i \in I)$. Then

$$b_x[T] = \inf_{i \in I} b_{x_i}[T_i] \leq \inf_{j \in J} b_{x_j}[T_j] \leq \inf_{j \in J} 1/j = 0.$$

Hence T cannot be a dynamics on S. □

So the **D**-product of the family $\{(S_i, D_i) : i \in I\}$ is defined as (S, D), where $S = \Pi\, S_i$, and $D =$ the subset of $\Pi\, D_i$ such that $T = (T_i : i \in I) \in D$ if and only if $\operatorname{dom} T_i = S_i \times \mathbb{R}$ for all but finitely many $i \in I$. Note that $I_S = (I_{S_i} : i \in I)$ (the trivial dynamics on S) is such that $\operatorname{dom} I_{S_i} = S_i \times \mathbb{R}$ for all $i \in I$, hence $I_S \in D$ always. Also, note that $D = \Pi\, D_i$ when I is finite.

It is interesting to note the resemblance of the definition of *product dynamics* to that of the *product topology*. Recall that in a product topological space $X = \Pi\, X_i$, sets of the form $A = \Pi\, A_i$, where each A_i is open in X_i and $A_i = X_i$ for all but finitely many i, form a base for the product topology.

Since **D** has products, it has a final object. The final object in **D** is $(1, \{I_1\})$, where 1 is a singleton set. For any dynamical system (S, D), the unique **D**-morphism $\phi: (S, D) \to (1, \{I_1\})$ is clearly the one sending all $x \in S$ to 1 and all $T \in D$ to I_1.

13. Coproducts

Coproducts of dynamical systems can also be defined. Let $\{(S_i, D_i) : i \in I\}$ be a family in **D**, let the coproduct phase space be $S = \amalg\, S_i = \cup(\{i\} \times S_i)$, and let the coproduct dynamics D be $\amalg\, D_i$. A dynamics $(i, T) \in D$ is defined by

$$(i, T)((j, x), t) = \begin{cases} (i, T(x, t)) & \text{if } j = i, \quad \text{for } a_x[T] < t < b_x[T] \\ (j, x) & \text{if } j \neq i, \quad \text{for } t \in \mathbb{R}. \end{cases}$$

So, roughly, (i, T) restricted to $\{i\} \times S_i$ is T on S_i and (i, T) restricted to $\amalg\, S_j : j \neq i$ is the trivial dynamics. Note we have $a_{(i,x)}[(i, T)] = a_x[T]$, $b_{(i,x)}[(i, T)] = b_x[T]$, and for $j \neq i$ $a_{(j,x)}[(i, T)] = -\infty$, $b_{(j,x)}[(i, T)] = +\infty$, i.e., $\operatorname{dom}(i, T) = (\{i\} \times \operatorname{dom} T) \bigcup \amalg (S_j \times \mathbb{R} : j \neq i)$.

The canonical injection maps $\iota_j : (S_j, D_j) \to (S, D)$ with $x \mapsto (j, x)$ and $T \mapsto (j, T)$ are such that

$$\iota_j(T(x, t)) = (j, T(x, t)) = (j, T)((j, x), t) = \iota_j T(\iota_j(x), t),$$

i.e., they are **D**-morphisms. One easily checks that this construction satisfies the universal property for categorical coproducts.

The initial object of **D** is $(\varnothing, \{I_\varnothing\})$. For any dynamical system (S, D), the unique **D**-morphism from $(\varnothing, \{I_\varnothing\})$ to (S, D) is the inclusion of $\varnothing$ in S with $I_\varnothing \mapsto I_S$.

E. Hierarchy of D-Morphisms and Image Factorizations

14. Monomorphisms

The hierarchy for **D**-monomorphisms is

split mono $\Rightarrow$ equalizer $\Leftrightarrow$ mono $\Leftrightarrow$ injection on S and D.

It is clear that a **D**-morphism that is injective on both S and D is mono. To show the converse, let $\phi:(S_1, D_1) \to (S_2, D_2)$ be a mono. Suppose there are distinct x and y in S_1 with $\phi x = \phi y$. Then ψ_1, $\psi_2:(S_1, D_1) \to (S_1, D_1)$, with ψ_1 sending all of S_1 to x, ψ_2 sending all of S_1 to y, and ψ_1, ψ_2 = identity on D_1, are **D**-morphisms with $\phi \circ \psi_1 = \phi \circ \psi_2$ but $\psi_1 \neq \psi_2$, a contradiction. So ϕ must be injective on S_1. Likewise suppose there are distinct T and T' in D_1 and $\phi T = \phi T'$. There are two cases to consider: either $T(x,t) = T'(x,t)$ whenever $(x,t) \in \operatorname{dom} T \bigcap \operatorname{dom} T'$ (hence $\operatorname{dom} T \neq \operatorname{dom} T'$), or there exists $(x_0, t_0) \in \operatorname{dom} T \bigcap \operatorname{dom} T'$ for which $T(x_0,t_0) \neq T'(x_0,t_0)$. In the former case defined U to be the restriction of T (and T') to $\operatorname{dom} T \bigcap \operatorname{dom} T'$, then ψ_1, $\psi_2:(S_1, \{U\}) \to (S_1, D_1)$, with ψ_1, ψ_2 = identity on S_1, $\psi_1 U = T$, $\psi_2 U = T'$, are **D**-morphisms with $\phi \circ \psi_1 = \phi \circ \psi_2$ but $\psi_1 \neq \psi_2$. In the latter case, let $S = \operatorname{dom} T_{t_0} \bigcap \operatorname{dom} T'_{t_0}$; then ψ_1, $\psi_2:(S, \{I_S\}) \to (S_1, D_1)$, with $\psi_1(x) = T(x,t_0)$, $\psi_2(x) = T'(x,t_0)$, $\psi_1 I_S = \psi_2 I_S = I_{S_1}$, are **D**-morphisms such that $\phi \circ \psi_1 = \phi \circ \psi_2$, because $(\phi \circ \psi_1)(x) = \phi(T(x,t_0)) = \phi T(\phi x, t_0) = \phi T'(\phi x, t_0) = \phi(T'(x,t_0)) = (\phi \circ \psi_2)(x)$. Note that $x_0 \in S$ is such that $\psi_1(x_0) = T(x_0,t_0) \neq T'(x_0,t_0) = \psi_2(x_0)$, hence $\psi_1 \neq \psi_2$. In both cases we are led to contradictions. So ϕ must be injective on D_1 as well. Thus a **D**-mono is injective on both the sets of states and dynamics.

To show that a mono $\phi:(S_1, D_1) \to (S_2, D_2)$ in **D** is an equalizer $\phi = \operatorname{eq}(\psi_1, \psi_2)$, ψ_1 and ψ_2 are defined as follows. First, let $S = \{0\} \times \phi(S_1) \bigcup \{1,2\} \times (S_2 \sim \phi(S_1))$—i.e., S is the disjoint union of one copy of $\phi(S_1)$ and two copies of $S_2 \sim \phi(S_1)$—and similarly let $D = \{0\} \times \phi(D_1) \bigcup \{1,2\} \times (D_2 \sim \phi(D_1))$. The members of D are defined in Table I. One can easily check that D is indeed a collection of dynamics on S [e.g., each $(0, T)$ is a dynamics on S because T induces a relative dynamics on $\phi(S_1)$ with the given relative bounds]. Now let ψ_1, $\psi_2:(S_2, D_2) \to (S, D)$ be

$$\psi_i(x) = \begin{cases} (0, x), & x \in \phi(S_1) \\ (i, x), & x \notin \phi(S_1) \end{cases} \qquad \psi_i T = \begin{cases} (0, T), & T \in \phi(D_1) \\ (i, T), & T \notin \phi(D_1) \end{cases}$$

Table I

Definition of Dynamics in D[a]

	$((0,x),t)$	$((1,x),t)$	$((2,x),t)$
$(0,T)$	$(0,T(x,t))$[b]	$(1,T(x,t))$[c]	$(2,T(x,t))$[c]
$(1,T)$	$(0,T(x,t))$, if $T(x,t)\in\phi(S_1)$; $(1,T(x,t))$, if $T(x,t)\notin\phi(S_1)$.[c]	$(0,T(x,t))$, if $T(x,t)\in\phi(S_1)$; $(1,T(x,t))$, if $T(x,t)\notin\phi(S_1)$.[c]	$(2,x)$[d]
$(2,T)$	$(0,T(x,t))$, if $T(x,t)\in\phi(S_1)$; $(2,T(x,t))$, if $T(x,t)\notin\phi(S_1)$.[c]	$(1,x)$[d]	$(0,T(x,t))$, if $T(x,t)\in\phi(S_1)$; $(2,T(x,t))$, if $T(x,t)\notin\phi(S_1)$.[c]

[a] The value of the dynamics i at the point j is given in the (row i, column j)-position. $T\in D_2$ and $x\in S_2$.

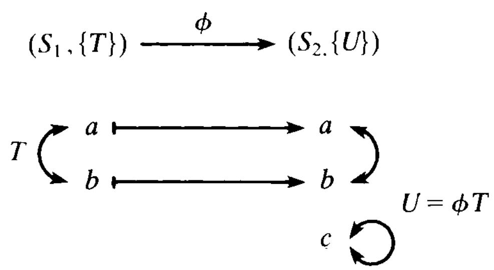
[b] For $a_y[U]<t<b_y[U]$, where $\phi y = x$, $\phi U = T$.

[c] For $a_x[T]<t<b_x[T]$.

[d] For $-\infty<t<+\infty$.

for $i=1,2$. One can check that ψ_1, ψ_2 are indeed **D**-morphisms, that $\phi(S_1)\subset S_2$ and $\phi(D_1)\subset D_2$ are the sets on which ψ_1 and ψ_2 agree, and $\psi_1\circ\phi=\psi_2\circ\phi$. Since $S_1\cong\phi(S_1)$ and $D_1\cong\phi(D_1)$, $\phi:(S_1,D_1)\to(S_2,D_2)$ is (up to isomorphism) the equalizer of ψ_1 and ψ_2. (It is clear, as in **S**, that the equalizer of a pair of morphisms in **D** is the inclusion map of the subsystem on which the morphisms agree.)

Finally, we show that an equalizer is not necessarily split mono. Let $S_1=\{a,b\}$ and $S_2=\{a,b,c\}$. Let $T:S_1\times\mathbb{R}\to S_1$ be defined by $T(a,0)=T(b,t)=a$, $T(b,0)=T(a,t)=b$ for $t\neq 0$. Let $U:S_2\times\mathbb{R}\to S_2$ be defined by $U(a,0)=U(b,t)=a$, $U(b,0)=U(a,t)=b$ for $t\neq 0$, and $U(c,t)=c$ for all t. It is easy to see that T and U are dynamics on S_1 and S_2, respectively. Now let $\phi:(S_1,\{T\})\to(S_2,\{U\})$ be $\phi(a)=a$, $\phi(b)=b$, and $\phi T=U$. Then ϕ is a **D**-morphism.



Any definition of ϕ^{-1} must have $\phi^{-1}(a) = a$, $\phi^{-1}(b) = b$, and $\phi^{-1}U = T$; but either $\phi^{-1}(c) = a$ or $\phi^{-1}(c) = b$ would be inconsistent: if $\phi^{-1}(c) = a$ then for $t \neq 0$, $\phi^{-1}(U(c,t)) = \phi^{-1}(c) = a$ but $T(\phi^{-1}(c), t) = T(a,t) = b$, hence the diagram

$$\begin{array}{ccc} S_2 \times \mathbb{R} & \xrightarrow{\phi^{-1} \times 1_{\mathbb{R}}} & S_1 \times \mathbb{R} \\ {\scriptstyle U}\downarrow & & \downarrow{\scriptstyle T} \\ S_2 & \xrightarrow[\phi^{-1}]{} & S_1 \end{array} \qquad \text{at} \qquad \begin{array}{ccc} (c,t) & \longmapsto & (a,t) \\ \downarrow & (t \neq 0) & \downarrow \\ c & \longmapsto & a \neq b \end{array}$$

is not commutative; similarly, for $\phi^{-1}(c) = b$. Thus ϕ does not split.

So we have established, in **D**,

$$\text{split mono} \Rightarrow \text{equalizer} \Leftrightarrow \text{mono} \Leftrightarrow \text{injection on } S \text{ and } D.$$

We can now define a **D**-*subobject*—a *dynamical subsystem*—(S_1, D_1) of (S_2, D_2) to be one such that there exists a mono (= equalizer = injection) ϕ from (S_1, D_1) to (S_2, D_2). Note that in this case, since $\phi(S_1) \cong S_1$, each $T \in D_1$ can be considered as the relative dynamics on S_1 induced by the dynamics ϕT on S_2.

15. Image Factorization Systems

Before we look at the hierarchy of **D**-epimorphisms, it would be useful to consider some more general concepts from category theory.

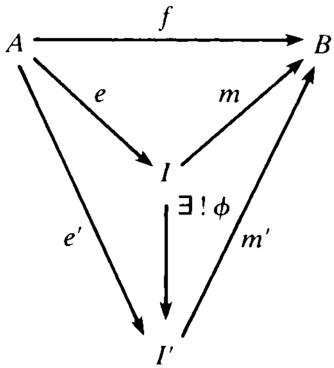
Definition. An *image factorization system* for a category **C** is a pair $(\mathscr{E}, \mathscr{M})$, where $\mathscr{E}$ and $\mathscr{M}$ are classes of **C**-morphisms such that

(i) $\mathscr{E}$ and $\mathscr{M}$ are subcategories of **C** [i.e., compositions of morphisms in $\mathscr{E}$ (respectively, $\mathscr{M}$) stay in $\mathscr{E}$ (respectively, $\mathscr{M}$)];

(ii) every element of $\mathscr{E}$ is an epi (possibly more; e.g., it could be a coequalizer) and every element of $\mathscr{M}$ is a mono (possibly more);

(iii) every isomorphism is in both $\mathscr{E}$ and $\mathscr{M}$;

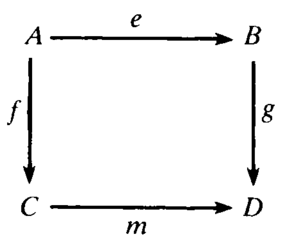
(iv) every morphism f in **C** has a unique (up to isomorphism) $\mathscr{E}$–$\mathscr{M}$ factorization; i.e., there exist $e \in \mathscr{E}$ and $m \in \mathscr{M}$ such that $f = m \circ e$, and whenever $e' \in \mathscr{E}$ and $m' \in \mathscr{M}$ satisfy $f = m' \circ e'$, there exists a unique isomorphism ϕ with $e' = \phi \circ e$ and $m' \circ \phi = m$.



If **C** has epi-equalizer factorizations, then $(\mathscr{E}, \mathscr{M})$ is an image factorization system if $\mathscr{E}$ = all epis and $\mathscr{M}$ = all equalizers. Dually, if **C** has coequalizer-mono factorizations, then $\mathscr{E}$ = all coequalizers and $\mathscr{M}$ = all monos form an image factorization system for **C**.

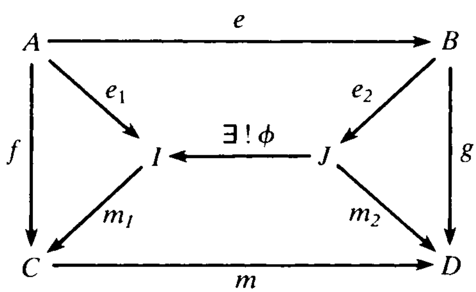
A consequence of (iv) in the definition is that if a morphism is in both $\mathscr{E}$ and $\mathscr{M}$ then it is an isomorphism; i.e., the converse of (iii) holds.

THE DIAGONAL FILL-IN LEMMA. Given the commutative square



with $e \in \mathscr{E}$ and $m \in \mathscr{M}$, there exists a unique $h: B \to C$ with $f = h \circ e$ and $g = m \circ h$.

Proof. Consider the diagram



Define $h = m_1 \circ \phi \circ e_2$. □

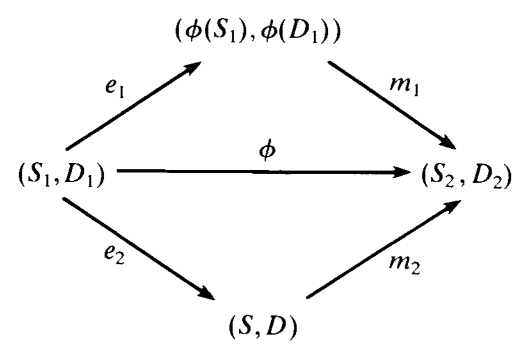
16. Epimorphisms

The hierarchy for **D**-epimorphisms is

$$\text{split epi} \Rightarrow \text{coequalizer} \Leftrightarrow \text{epi} \Leftrightarrow \text{surjection on } S \text{ and } D.$$

It is clear that a **D**-morphism that is surjective onto S and D is epi. For the converse, let $\phi:(S_1, D_1) \to (S_2, D_2)$ be an epi. Suppose there exists $U \in D_2 \sim \phi(D_1)$. Let $S = S_2 \bigcup \{a, b\}$ where $a, b \notin S_2$, and $D = \{0, 1\} \times D_2$. $(0, T) \in D$ is defined on S as T on S_2 and identity on $\{a, b\}$. $(1, T) \in D$ is defined on S as T on S_2 and it "interchanges" a and b (i.e., $(1, T)(a, 0) = (1, T)(b, t) = a$ and $(1, T)(b, 0) = (1, T)(a, t) = b$ for $t \neq 0$). Then (S, D) is a **D**-object. Let $\psi_1:(S_2, D_2) \to (S, D)$ be defined by sending each $x \in S_2$ to itself and sending each $T \in D_2$ to $(0, T)$; and let $\psi_2:(S_2, D_2) \to (S, D)$ be defined by sending each $x \in S_2$ to itself and sending each $T \in D_2$ to $(0, T)$ *except* $\psi_2 \mathrm{U} = (1, \mathrm{U})$. Then one sees that ψ_1 and ψ_2 are **D**-morphisms with $\psi_1 \circ \phi = \psi_2 \circ \phi$ but $\psi_1 \neq \psi_2$, a contradiction. So ϕ must be onto D. If $S_2 \sim \phi(S_1)$ is nonempty, consider $\psi_1, \psi_2:(S_2, D_2) \to (\{0, 1\}, \{I_{\{0,1\}}\})$ with $\psi_1 = \chi_{\phi(S_1)}$ on S_2, $\psi_2 = \chi_{S_2}$ on S_2 and $\psi_1 T = \psi_2 T = I_{\{0,1\}}$ for all $T \in D_2$. Then since T induces a relative dynamics on $\phi(S_1)$ (because $T \in D_2 = \phi(D_1)$), ψ_1 is a **D**-morphism; ψ_2 is trivially a **D**-morphism. And $\psi_1 \circ \phi = \psi_2 \circ \phi$ with $\psi_1 \neq \psi_2$, a contradiction. So ϕ must also be onto S_2. Thus an epi is onto both S and D.

To show that an epi in **D** is a coequalizer, we shall use the diagonal fill-in lemma from the previous section. Let $\phi:(S_1, D_1) \to (S_2, D_2)$ be a **D**-morphism, let $\phi = m_1 \circ e_1$ be its epi-equalizer factorization, and let $\phi = m_2 \circ e_2$ be its coequalizer-mono factorization. Let $\mathscr{E}$ = all **D**-epis and $\mathscr{M}$ = all **D**-equalizers, hence $(\mathscr{E}, \mathscr{M})$ is an image factorization system for **D**. We have the commutative diagram



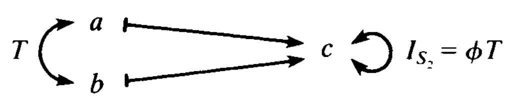
[It is clear that the "image object" in the epi-equalizer factorization is $(\phi(S_1), \phi(D_1))$.]

Now consider the pair (e_1, m_2). $e_1 \in \mathscr{E}$, and m_2 is a **D**-mono hence a **D**-equalizer (see Section IV.14), so $m_2 \in \mathscr{M}$. Thus by the diagonal fill-in lemma there exists a unique $g:(\phi(S_1), \phi(D_1)) \to (S, D)$ with $e_2 = g \circ e_1$ and $m_1 = m_2 \circ g$. On the other hand, the pair (e_2, m_1) is such that e_2 is a coequalizer and hence an epi (general hierarchy), so $e_2 \in \mathscr{E}$, and $m_1 \in \mathscr{M}$. Thus there exists a unique $h:(S, D) \to (\phi(S_1), \phi(D_1))$ with $e_1 = h \circ e_2$ and $m_2 = m_1 \circ h$. Now $e_2 = g \circ e_1 = g \circ h \circ e_2$ and $e_1 = h \circ e_2 = h \circ g \circ e_1$, so $g \circ h = 1_{(S,D)}$ and $h \circ g = 1_{(\phi(S_1), \phi(D_1))}$ (because e_1 and e_2 are epi), whence (S, D) is isomorphic to $(\phi(S_1), \phi(D_1))$ in **D**.

This shows that the epi-equalizer and coequalizer-mono factorizations in **D** are the same and so in particular an epi is a coequalizer.

The following simple example shows that a **D**-coequalizer is not necessarily a split epi. Let $S_1 = \{a, b\}$ and $S_2 = \{c\}$. Let T be the dynamics on S_1 defined by $T(a,0) = T(b,t) = a$, $T(b,0) = T(a,t) = b$ for $t \neq 0$. Then $\phi:(S_1, \{T\}) \to (S_2, \{I_{S_2}\})$ with $\phi(a) = \phi(b) = c$ and $\phi T = I_{S_2}$ is clearly epi, hence a coequalizer.

$$(S_1, \{T\}) \xrightarrow{\ \phi\ } (S_2, \{I_{S_2}\})$$



If ϕ^{-1} exists, it has to have $\phi^{-1}(c) = a$ or b. But one easily checks that neither case would yield a **D**-morphism.

So we have, in **D**,

$$\text{split epi} \Rightarrow \text{coequalizer} \Leftrightarrow \text{epi} \Leftrightarrow \text{surjection on } S \text{ and } D.$$

And, quite different from **S**, there is only one image factorization in **D**—epi-equalizer = coequalizer-mono = surjection-injection.

F. Dynamics and Equivalence Relations

17. Compatibility

DEFINITION. Let T be a dynamics on S. T is *compatible* with an equivalence relation R on S if xRy implies

(i) $T(x,t)$ is defined if and only if $T(y,t)$ is defined, and
(ii) $T(x,t) R T(y,t)$.

When dom $T = S \times \mathbb{R}$, condition (i) is of course superfluous and the compatibility condition reduces to the usual one for the usual definition of dynamics (cf. Rosen (1978)).

18. Quotient Dynamics

If T is compatible with the equivalence relation R on S, then T induces a dynamics T' on the set of 'reduced states' $S' = S/R$. For any $(x)_R \in S'$ and $t \in \mathbb{R}$, let $T'((x)_R, t) = (T(x,t))_R \in S'$ whenever $T(x,t)$ is defined, and leave $T'((x)_R, t)$ undefined otherwise.

To show that T' is a well-defined function, we have to show that it is independent of the equivalence class representative x. But for xRy, $T(x,t)$ is defined if and only if $T(y,t)$ is defined, and $T(x,t)RT(y,t)$, whence $(T(x,t))_R = (T(y,t))_R$. So T' is indeed independent of the choice of equivalence class representatives. Further, condition IV.17(i) actually implies that when xRy, $a_x[T] = a_y[T]$ and $b_x[T] = b_y[T]$, so we obtain the well-defined bounds for T' at $(x)_R$ as $a_{(x)_R}[T'] = a_x[T]$ and $b_{(x)_R}[T'] = b_x[T]$.

The initial value property $T'((x)_R, 0) = (T(x,0))_R = (x)_R$ clearly holds for T'. To show that T' is indeed a dynamics, it remains to verify the group property of IV.3(iii). Suppose $T'((x)_R, t_1) = (T(x,t_1))_R$ is defined (thus $T(x,t_1)$ is defined, in particular); then if $T'(T'((x)_R, t_1), t_2) = T'((T(x,t_1))_R, t_2)$ is defined, it is equal to $(T(T(x,t_1),t_2))_R = (T(x, t_1 + t_2))_R = T'((x)_R, t_1 + t_2)$, using the group property for T. Similarly, when both $T'((x)_R, t_1)$ and $T'((x)_R, t_1 + t_2)$ are defined, we obtain the same equality.

The dynamics T' is called the *quotient dynamics* on S' induced by T.

REMARK. Note that the equation $T'((x)_R, t) = (T(x,t))_R$ states that the diagram

$$\begin{array}{ccc} \text{dom } T & \xrightarrow{\chi \times 1_{\mathbb{R}}} & \text{dom } T' \\ {\scriptstyle T}\big\downarrow & & \big\downarrow{\scriptstyle T'} \\ S & \xrightarrow[\chi]{} & S' \end{array}$$

commutes (where $\chi: S \to S'$ is the quotient map $x \mapsto (x)_R$). So $\chi: (S, \{T\}) \to (S', \{T'\})$ sending each $x \in S$ to $(x)_R \in S'$ and sending T to T' is a **D**-morphism (in fact, a **D**-epimorphism).

19. Bijections and Equivalence Relations

There are two complementary questions on the connection between dynamics on S and equivalence relations on S that can be asked:

(1) Given a dynamics T on S, how can we characterize those equivalence relations R on S with which T is compatible?

(2) Given an equivalence relation R on S, how can we characterize those dynamics on S that are compatible with R?

The simpler case when one considers bijections ('automorphisms') $T: S \to S$ instead of dynamics was discussed in Rosen (1978) [in which a bijection $T: S \to S$ is defined to be compatible with an equivalence relation R on S if sRs' implies $(Ts)R(Ts')$]. It was shown that the set of all equivalence relations with which a given bijection $T: S \to S$ is compatible forms a sublattice of the lattice $\mathscr{R}(S)$ of all equivalence relations on S and that the set of all bijections compatible with a given equivalence relation forms a submonoid of the group $\mathscr{A}(S)$ of all bijections on S; further, the set of all bijections compatible with R, and whose inverses are also compatible with R, forms a subgroup of $\mathscr{A}(S)$. [Note that when both T and T^{-1} are compatible with R, sRs' iff $(Ts)R(Ts')$.]

Let us first expand on these ideas.

20. Galois Theory

DEFINITION. A subset $\mathscr{A}$ of $\mathscr{A}(S)$ is *compatible* with a family $\mathscr{R}$ of equivalence relations on S if for each $T \in \mathscr{A}$, for each $R \in \mathscr{R}$, and for all $s, s' \in S$, sRs' implies $(Ts)R(Ts')$.

We shall now try to answer these two questions:

(1*) Given a subgroup G of $\mathscr{A}(S)$, what are the equivalence relations on S with which G is compatible?

(2*) Given a sublattice L of $\mathscr{R}(S)$, what are the bijections on S such that they and their inverses are compatible with L?

The same argument used to obtain the results of Section IV.19 can easily be generalized to answer these questions. The answer to (1*) is a sublattice of $\mathscr{R}(S)$ which contains the equality relation $=$ and the "universal" relation S^2 defined by sS^2s' iff $s, s' \in S$. Since every bijection is compatible with $=$ and S^2, it is without loss of generality to (and hence we shall) only consider sublattices that contain both of these relations. The answer to (2*) is a subgroup of $\mathscr{A}(S)$. In fact, this correspondence between the set of all sublattices of $\mathscr{R}(S)$ and the set of all subgroups of $\mathscr{A}(S)$ turns out to be bijective and order (*qua* substructures)-inverting. We can represent this situation by the following self-explanatory diagram:

$$\begin{array}{ccccccc}
\mathscr{R}(S) & \longleftarrow & L_1 & \longleftarrow & L_2 & \longleftarrow & \{=, S^2\} \\
\updownarrow & & \updownarrow & & \updownarrow & & \updownarrow \\
1 & \rightarrowtail & G_1 & \rightarrowtail & G_2 & \rightarrowtail & \mathscr{A}(S)
\end{array}$$

Further, we can define that an equivalence relation R' on S is *conjugate* to another equivalence relation R on S if there is a bijection $T: S \to S$ such that $sR's'$ if and only if $(Ts)R(Ts')$, in which case we shall write $R' = R_T$. Conjugacy is equivalent to the existence of an isomorphism between the collections of equivalence classes S/R and S/R', i.e., there is a bijection between the equivalence classes in S/R and S/R' and that corresponding equivalence classes contain the same number of S-elements.

Let L be a sublattice of $\mathscr{R}(S)$, then we can define a *conjugate* of L to be $L_T = \{R_T: R \in L\}$ for some $T \in \mathscr{A}(S)$. With these definitions we have the following interesting lemma.

LEMMA. If (L, G) is a pair of corresponding sublattice-subgroup with respect to compatibility, then the conjugate $(L_T, T^{-1}GT)$ is also a pair of corresponding sublattice-subgroup with respect to compatibility.

Proof. Let $R \in L$ and $U \in G$, then for $s, s' \in S$,

$$
\begin{aligned}
sR_Ts' \quad &\text{iff} \quad (Ts)R(Ts') && [\text{definition of } R_T] \\
&\text{iff} \quad (UTs)R(UTs') && [(R, U) \in (L, G)] \\
&\text{iff} \quad (TT^{-1}UTs)R(TT^{-1}UTs') && [TT^{-1} = 1_S] \\
&\text{iff} \quad (T^{-1}UTs)R_T(T^{-1}UTs') && [\text{definition of } R_T]
\end{aligned}
$$

Thus $T^{-1}UT$ and $(T^{-1}UT)^{-1}$ are both compatible with R_T. □

The preceding results bear a striking resemblence to the Galois theory of field extensions and automorphism groups!

21. Discrete Dynamical Systems

Before going further, let us take a digression in the following direction. Suppose instead of considering time as the continuum of real numbers we consider time as being composed of a succession of "elementary steps." Then we can define a *discrete dynamics* on a set S to be a mapping T from $S \times \mathbb{Z}$ to S such that

(i) $T(x, 0) = x$ for all $x \in S$, and
(ii) $T(T(x, t_1), t_2) = T(x, t_1 + t_2)$ for all $x \in S$ and all $t_1, t_2 \in \mathbb{Z}$.

For simplicity we have assumed that dom T is all of $S \times \mathbb{Z}$ instead of a subset as in Definition IV.3 of a "continual" dynamics.

It is clear that for each $n \in \mathbb{Z}$, each $T(\cdot, n)$ is a bijection from S to S and that $T(\cdot, n) = T(\cdot, 1)^n$. Thus, alternatively, we can define a discrete dynamics to be the cyclic subgroup of $\mathscr{A}(S)$ generated by a bijection $T \in \mathscr{A}(S)$, i.e., a discrete dynamics is $\langle T \rangle = \{T^n : n \in \mathbb{Z}\} \subset \mathscr{A}(S)$. With this terminology the powers of

T in $\langle T \rangle$ become interpretable as *instants of logical time*, and the transition $x \mapsto Tx$, or $T^{n-1}x \mapsto T^n x$ in general, is an *elementary step* of the dynamics. Further, with appropriate modifications, all the discussions on the category **D** still go through with $\mathbb{R}$ replaced by $\mathbb{Z}$ and we would have the category of *discrete* dynamical systems instead.

This situation is, of course, closely related to the specification of a continual dynamical system as a one-parameter group of bijections $T = \{T_t : t \in \mathbb{R}\}$ on a phase space S. (Let us again restrict to the case when $\operatorname{dom} T = S \times \mathbb{R}$, i.e., when $\operatorname{dom} T_t = S$ for all t.) For any real number r, we can consider the cyclic subgroup generated by T_r, then $\langle T_r \rangle = \{T_r^n = T_{nr} : n \in \mathbb{Z}\}$ defines a discrete dynamics on S. This method of "discretization" is used, for example, in obtaining numerical solutions of differential equations. Note, however, that this procedure only goes one way: We can obtain a discrete time from a continuous time by choosing the size of an elementary step $t = r$, but starting from a discrete dynamics $\langle T \rangle$, in general we cannot embed it into a continuous one-parameter group of bijections. A further discussion of this aspect can be found in Rosen (1981). We shall come back to the role of time in dynamics at the end of Section V.

22. Discrete Dynamics and Equivalence Relations

Now let us try to answer the two questions in Section IV.19 for discrete dynamics.

Given a discrete dynamics $\langle T \rangle = \{T^n : n \in \mathbb{Z}\}$, $T \in \mathscr{A}(S)$, since $\langle T \rangle$ is in particular a (cyclic) subgroup of $\mathscr{A}(S)$, the results of Section IV.20 tell us that the set of all equivalence relations with which $\langle T \rangle$ is compatible forms a sublattice of $\mathscr{R}(S)$. It is easy to see that $\langle T \rangle$ is compatible with an equivalence relation R on S if and only if both T and T^{-1} are compatible with R, so we only need to check the compatibility of the generator and its inverse with R.

Since given an equivalence relation R on S, the set of all bijections on S compatible with R and whose inverses are also compatible with R forms a subgroup G of $\mathscr{A}(S)$, the set of all discrete dynamics on S compatible with R is the collection of all cyclic subgroups of G, i.e., the collection of all homomorphic images of $(\mathbb{Z}, +)$ in G.

23. Partial Answers

Finally, we shall try to answer, at least partially, the two questions in Section IV.19. We shall only consider the dynamics T for which $\operatorname{dom} T = S \times \mathbb{R}$.

Given a dynamics T on S, the collection of translations $\{T_t : t \in \mathbb{R}\}$ is a subgroup of $\mathscr{A}(S)$; thus the set of all equivalence relations with which T is compatible is a sublattice of $\mathscr{R}(S)$.

The second question is more difficult. Given an equivalence relation R on S, we obtain a subgroup G of $\mathscr{A}(S)$ of all bijections on S compatible with R. Now we have to find subgroups of G that are (isomorphic to) continuous one-parameter groups indexed by $\mathbb{R}$. Whereas for the discrete dynamics case we can get the set $\{\{T^n : n \in \mathbb{Z}\} : T \in G\}$ quite easily, here there is no trivial way to look for homomorphic images of $(\mathbb{R}, +)$ in G. Note, however, that there is at least one dynamics compatible with R, namely, the trivial dynamics $I_S = \{1_S : t \in \mathbb{R}\} \subset G$.

V. TOPOLOGICAL DYNAMICS

That which gives things their suchness
Cannot be delimited by things.
So when we speak of "limits" we remained confined
To limited things.

Chuang Tse

In this section the study of dynamical systems is continued. First, in Subsections A and B, we shall consider the topology on the phase space induced by a dynamics; then we shall consider an arbitrary topology on the phase space of a dynamics such that the dynamics is a continuous map. The development in these sections is adapted from Hájek (1968). Next, via topological constructions, the relations between observables and dynamics on the same set are analysed in Subsection C. Finally, in Subsection D, the role of time in dynamics is discussed.

A. *T*-Topology

Let T be a dynamics on a set S.

DEFINITION. A subset $G \subset S$ is *T-open* if for every $x \in G$ there exists an $\varepsilon > 0$ such that $T(x, (-\varepsilon, \varepsilon)) = \{T(x, t): -\varepsilon < t < \varepsilon\} \subset G$.

The collection of T-open sets clearly forms a topology on S with the collection $\{T(x, (-\varepsilon, \varepsilon)) : x \in S, \varepsilon > 0\}$ as a base. This topology is called the *T-topology* on S.

1. Properties

Some immediate consequences of the definition are:

(i) A net $\{x_i\}$ is *T-convergent* to x in S if and only if for large i, $x_i = T(x, t_i)$ with $t_i \to 0$ in $\mathbb{R}$. This also characterizes *T-closed* sets.

(ii) S is *T-locally compact.*

(iii) S is *T-locally pathwise connected.*

(iv) A subset of S is *T-invariant* if and only if it is *T-clopen* (closed and open). So all trajectories of T are T-clopen (see Section IV.7).

(v) S is *T-connected* if and only if it consists of a single trajectory (or if S is empty).

(vi) If S is *T-compact*, then it consists of a finite set of trajectories.

LEMMA 1. $\operatorname{dom} T \subset S \times \mathbb{R}$ is T-open. The map $T: \operatorname{dom} T \to S$ is T-continuous. (The topologies on the various sets are, naturally, the T-topology for S, the usual topology for $\mathbb{R}$, the product topology for $S \times \mathbb{R}$, and the subspace topology for $\operatorname{dom} T \subset S \times \mathbb{R}$.)

Proof. It suffices to show that the inverse image under T of the basic T-open set $G = T(x, (-\varepsilon, \varepsilon))$ is open in $S \times \mathbb{R}$. Now if $(x', t') \in T^{-1}(G)$, then $T(x', t') = T(x, t)$ with $|t| < \varepsilon$. Let

$$\delta = \tfrac{1}{2}(\varepsilon - |t|)$$

and consider $E = T(x', (-\delta, \delta)) \times (t' - \delta, t' + \delta)$. E is then clearly an open neighbourhood of (x', t') in $S \times \mathbb{R}$ with $T(E) = T(T(x', t'), (-2\delta, 2\delta)) = T(T(x, t), (|t| - \varepsilon, \varepsilon - |t|)) = T(x, (t + |t| - \varepsilon, t - |t| + \varepsilon)) \subset T(x, (-\varepsilon, \varepsilon)) = G$. Thus $T^{-1}(G)$ is open in $S \times \mathbb{R}$. □

LEMMA 2. The T-topology is the finest among topologies on S rendering all solutions $y_x: (a_x, b_x) \to S$ continuous.

Proof. It follows from Lemma 1 that all y_x are T-continuous. On the other hand, let τ be a topology on S in which all y_x are continuous. Let $G \in \tau$ and $x \in G$. Then $y_x(0) = x \in G$ so $T(x, (-\varepsilon, \varepsilon)) = y_x(-\varepsilon, \varepsilon) \subset G$ for small $\varepsilon > 0$. Thus G is open in the T-topology, which is therefore finer than τ. □

THEOREM 3. The T-topology is the finest among topologies on S rendering $T: \operatorname{dom} T \to S$ continuous. □

COROLLARY. The bounds a_x and b_x, considered as maps from S to $\bar{\mathbb{R}}$, are T-continuous.

Proof. Theorem 3 and Lemma IV.6. □

LEMMA 4. Let $A \subset S$. Then T induces a relative dynamics T' on A if and only if A is T-open in S. In the positive case, the T'-topology on A coincides with the subspace T-topology of A in S.

Proof. $\Rightarrow$ If $x \in A$, then by the definition of dynamics, $T'(x, t) \in A$ for small $|t|$. Since T' is the restriction of T, we have $T(x, t) \in A$ for small $|t|$, so A is T-open.

$\Leftarrow$ Suppose A is T-open. Let $x \in A$. For $t \geq 0$ define $T'(x,t) = T(x,t)$ if $T(x,t') \in A$ for all $t' \in [0,t]$. Similarly, for $t \leq 0$. Otherwise leave $T'(x,t)$ undefined. Then clearly T' is the relative dynamics on A induced by T. The bounds $a_x[T']$ and $b_x[T']$ are determined by

$$a_x[T'] = \inf\{t : T(x,t') \in A \text{ for } t \leq t' \leq 0\},$$

$$b_x[T'] = \sup\{t : T(x,t') \in A \text{ for } 0 \leq t' \leq t\}.$$

It is clear that the two topologies on A coincide. $\square$

Compare IV.7, property (iv), and this lemma:

A is T-invariant $\xleftrightarrow{\text{(iv)}}$ A is T-clopen

IV.7 $\Downarrow$ $\Downarrow$

T induces a relative dynamics on A $\xleftrightarrow[\text{Lemma 4}]{}$ A *is* T-open

LEMMA 5. Let $\phi:(2_1,D_1) \to (S_2,D_2)$ be a **D**-morphism. Then for each $T \in D_1$, $\phi: S_1 \to S_2$ is a continuous function where S_1 and S_2 are endowed with the T- and ϕT-topology, respectively.

Proof. It follows from Lemma 2 that $\phi: S_1 \to S_2$ is continuous if and only if $\phi \circ y_x: \mathbb{R} \to S_1 \to S_2$ is continuous for each solution y_x of T. By definition of a **D**-morphism $\phi(T(x,t)) = \phi T(\phi x, t)$, hence $\phi(y_x(t)) = y'_{\phi x}(t)$ (where $y'_{x'}$ denotes solutions of ϕT), and $a_{\phi x}[\phi T] \leq a_x[T] < 0 < b_x[T] \leq b_{\phi x}[\phi T]$ (see Definition IV.9). Thus $\operatorname{dom}(\phi \circ y_x) = (a_x[T], b_x[T])$ is an open (in $\mathbb{R}$) subset of dom $y'_{\phi x} = (a_{\phi x}[\phi T], b_{\phi x}[\phi T])$, and so $\phi \circ y_x$ is continuous because $y'_{\phi x}$ is. $\square$

B. Continuous Dynamics

2. *Definition and Consequences*

DEFINITION. Let T be a dynamics on S and τ be a topology on S. T is a *continuous dynamics* on (S,τ) (or simply "on S" if τ is clear or immaterial) if

(i) T: dom $T \to S$ is continuous, and
(ii) dom T is open in $S \times \mathbb{R}$

(where, of course, $S \times \mathbb{R}$ has the product topology of τ on S and the usual topology on $\mathbb{R}$).

Note condition (i) implies that τ is coarser than the T-topology on S [see Theorem V.1(3)]. If dom $T = S \times \mathbb{R}$, then (ii) is automatically satisfied so a continuous dynamics in this case is simply a dynamics that is continuous.

Conditions (i) and (ii) are equivalent to each of the following:

(a) If the net $\{x_i\}$ converges to x in S, the net $\{t_i\}$ converges to t in $\mathbb{R}$, and $T(x,t)$ is defined, then $T(x_i, t_i) \to T(x,t)$.

(b) The inverse image under T of an open set in S is open in $S \times \mathbb{R}$.

EXAMPLES. (1) T is a continuous dynamics on S with the T-topology [Lemma V.1(1)].

(2) T is a continuous dynamics on S with the discrete topology if and only if $T = I_S$ = the trivial dynamics on S. In fact, the discrete topology is the I_S-topology on S.

(3) The trivial dynamics on S is a continuous dynamics on S with any topology.

(4) T is a continuous dynamics on S with the indiscrete topology $(= \{S, \varnothing\})$ if and only if dom $T = S \times \mathbb{R}$.

(5) (See Section IV.4) The dynamics associated with an autonomous differential equation $dx/dt = f(x)$ on an open subset $S \subset \mathbb{R}^k$, where $f: S \to \mathbb{R}^k$ is as in Section IV.4, is a continuous dynamics on S with the usual topology. Note that condition (i) corresponds to continuous dependence on initial data.

LEMMA 1. *If T is a continuous dynamics on S, then the bounds a_x and b_x are, respectively, upper and lower semicontinuous maps $S \to \mathbb{R}$; i.e., $a_x \geq \limsup_{y \to x} a_y$ and $b_x \leq \liminf_{y \to x} b_y$.*

Proof. Let $x \in S$ and $t \in (a_x, 0]$. Then $(x,t) \in \operatorname{dom} T$, so there is a neighbourhood N around x such that for $y \in N$, $(y,t) \in \operatorname{dom} T$ (since dom T is open in $S \times \mathbb{R}$). Hence $t \in (a_y, 0]$ for these y, and so $t \geq \limsup_{y \to x} a_y$. Now take $t \to a_x$.

Similarly for b_x. □

COROLLARY. *If T is a continuous dynamics on S, then for any $t \in \mathbb{R}$, the sets $\{x \in S : a_x < t\}$ and $\{x \in S : t < b_x\}$ are open in S, so that $\{x \in S : a_x = -\infty\}$ and $\{x \in S : b_x = +\infty\}$ are G_δ's in S.* □

LEMMA 2. *If T is a continuous dynamics on S, then each solution y_x is continuous, and each translation T_t is continuous with domain open in S.* □

LEMMA 3. *If T is a continuous dynamics on (S, τ) and $A \subset S$ is either τ-open or T-invariant, then the relative dynamics on A induced by T is a continuous dynamics on A with the subspace τ-topology.*

Proof. It follows from Theorem V.1(3) and property V.1(iv) that in either case A is T-open, so Lemma V.1(4) gives this result. □

REMARK. Lemmata V.1(4) and V.2(3) give connections between continuity and **D**-subsystems. Lemma V.1(5) gives connections between continuity and **D**-morphisms. With Theorem V.1(3), this says that when $\phi:(S_1, D_1) \to (S_2, D_2)$ is a **D**-morphism, and if ϕT is a continuous dynamics on S_2 with some topology τ, then $\phi: S_1 \to S_2$ is continuous where S_1 has the T-topology and S_2 has the topology τ. Note, however, this does *not* imply that if T is a continuous dynamics on (S_1, τ'), then $\phi: S_1 \to S_2$ is continuous with respect to τ' and the ϕT-topology; the implication only goes one way.

Now how about the connections between continuity and **D**-products, **D**-coproducts, and quotient dynamics? The answers are given in the next few theorems.

THEOREM 4. For each $i \in I$ let T_i be a continuous dynamics on (S_i, τ_i). If the product dynamics $T = (T_i : i \in I)$ on $S = \Pi S_i$ exists in **D**, then T is a continuous dynamics on (S, τ) where τ is the product topology of the τ_i's.

Proof. By Theorem IV.12 there is a finite set $J \subset 1$ such that $\operatorname{dom} T_i = S_i \times \mathbb{R}$ for $i \notin J$.

To show that T is a continuous dynamics on (S, τ), it is necessary and sufficient to show that the inverse image of a τ-open set $G \subset S$ under T is open in $S \times \mathbb{R}$ [see condition (b)]. It suffices to let G be a subbasic open set $G = G_k \times \Pi_{i \neq k} S_i$ for some $k \in I$, G_k open in S_k. Also, without loss of generality assume $k \notin J$.

Let $x = (x_i : i \in I) \in S$, $t \in \mathbb{R}$, and $T(x,t) \in G$. Then in particular $T_k(x_k, t) \in G_k$ so condition (b) applied to T_k yields a τ_k-open $H_k \subset S_k$, open $E_k \subset \mathbb{R}$, with $(x_k, t) \in H_k \times E_k \subset T_k^{-1}(G_k) \subset \operatorname{dom} T_k$. Doing the same thing for each $i \in J$, we obtain two families of open sets $\{H_i : i \in J\}$ and $\{E_i : i \in J\}$. Let

$$H = H_k \times \prod_{i \in J} H_i \times \prod_{\substack{i \notin J \\ i \neq k}} S_i \subset S,$$

$$E = E_k \bigcap \bigcap_{i \in J} E_i \subset \mathbb{R}.$$

Then $H \times E$ is open in $S \times \mathbb{R}$ and $(x, t) \in H \times E \subset T^{-1}(G) \subset \operatorname{dom} T$, hence $T^{-1}(G)$ is open. □

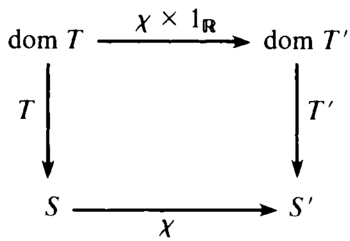
THEOREM 5. Let T_i be a continuous dynamics on (S_i, τ_i) for each $i \in I$. Then each of the coproduct dynamics (i, T_i) in **D** is a continuous dynamics on $S = \amalg S_i$ with the direct sum topology.

Proof. (See Section IV.13) The proof is immediate; recall that a set $G \subset S$ is open relative to the direct sum topology if and only if for each $i \in I$, $G \cap S_i$ is τ_i-open. □

THEOREM 6. Let T be a continuous dynamics on (S, τ) and let T be compatible with an equivalence relation R on S. Let T' be the quotient dynamics induced by T on $S' = S/R$, and let τ' be the quotient topology on S'. Then T' is a continuous dynamics on (S', τ').

Proof. (See Section IV.18) The equivalence relation R on S induces an equivalence relation R on $S \times \mathbb{R}$ via $(x, t)R(x', t')$ if and only if xRx' and $t = t'$. Then $(S \times \mathbb{R})/R \cong S' \times \mathbb{R}$. Also, since the bounds are such that $a_{(x)_R}[T'] = a_x[T]$ and $b_{(x)_R}[T'] = b_x[T]$, we have dom $T/R \cong$ dom T', and the topology on dom T' is in fact the quotient of the topology on dom T modulo R.

Now consider the commutative diagram where $\chi: S \to S'$ is the natural projection, and hence $\chi \times 1_{\mathbb{R}}$: dom $T \to$ dom T' = dom T/R is also the corresponding natural projection:



Let $G \subset S'$ be τ'-open. Since both T and χ are continuous, $(\chi \circ T)^{-1}(G) \subset$ dom T is open (in the product topology of $S \times \mathbb{R}$). But $(\chi \circ T)^{-1}(G) = (\chi \times 1_{\mathbb{R}})^{-1} \circ (T')^{-1}(G)$, so by the definition of the quotient topology on dom T', $(T')^{-1}(G)$ must be open (in the topology of $S' \times \mathbb{R}$). Thus T' is a continuous dynamics on (S', τ'). □

C. Observables and Induced Topologies

3. Observables Revisited

Let f be an observable on the set of states S, i.e., $f: S \to \mathbb{R}$. In the construction of the category **S**, we could have used equivalence classes of functions in $\mathbb{R}^S/\sim$ as observables (see Section III.2) because there the only relevant property of f was the equivalence relation R_f it imposed on S. But such a definition of **S**-objects apparently leads to difficulties in some categorical constructions, and creates problems in topological considerations.

But we should recall the comments in Section III.7 and note that for $f, g \in \mathbb{R}^S$ and $f \sim g$ (i.e., $R_f = R_g$), $(S, \{f\})$ and $(S, \{g\})$ are isomorphic in **S**, and because all constructions in a category are only "up to isomorphism," these two **S**-objects (and any constructions with one or the other) are "indistinguishable" in **S**.

4. *Induced Topology*

What properties of $f \in \mathbb{R}^S$ are hidden if we consider $(f)_\sim \in \mathbb{R}^S/\sim$ instead? We see that since $\mathbb{R}$ is a topological space (with the usual topology), $f: S \to \mathbb{R}$ can induce a topology on S, called the *f-topology*, as follows. A subset of S is *f-open* (respectively, *f-closed*) if and only if it is the inverse image under f of an open (respectively, closed) subset of $\mathbb{R}$. The f-topology is the coarsest topology on S that renders f continuous.

For any constant function f [i.e., for $f \in (0)_\sim \in \mathbb{R}^S/\sim$], the f-topology is the indiscrete topology $\{\emptyset, S\}$ on S. And in general if the range of f is a finite set, then any representative of the class $(f)_\sim$ induces the same f-topology on S, and so there is a unique $(f)_\sim$-topology. But if the range of f is infinite, then it is possible for $f(S)$ to have limit points, in which case different class representatives of $(f)_\sim$ may induce different topologies; hence one has to consider $f \in \mathbb{R}^S$ and not $(f)_\sim \in \mathbb{R}^S/\sim$ in topological considerations.

Further, since $\mathbb{R}$ is a metric space, a "distance function" on S can be defined using f. Namely, for $x, y \in S$, define $d_f(x, y) = |f(x) - f(y)|$. It is clear that $d_f(x, x) = 0$, that $0 \leq d_f(x, y) = d_f(y, x) < +\infty$, and that d_f satisfies the triangle inequality. But $d_f(x, y) = 0$ only means $f(x) = f(y)$ and not necessarily $x = y$, so d_f is a *pseudometric* on S, and d_f is a metric if and only if f is injective. Obviously, the f-topology on S is the pseudometric topology generated by d_f, and the quotient f-topology on S/R_f is the induced metric topology. Note that if $x \in G \subset S$ where G is f-open and $d_f(x, y) = 0$, then $y \in G$.

Historically, the ideas of limit and continuity appeared very early in mathematics, notably in geometry, and their role has steadily increased with the development of analysis and its applications to the experimental sciences, since these ideas are closely related to those of *experimental determination* and *approximation*. But since most experimental determinations are *measurements*, i.e., determinations of one or more numbers, it is hardly surprising that the notions of limit and continuity in mathematics were featured at first only in the theory of real numbers and its outgrowths and fields of application. So in a sense topology has its roots in the process of measurement (i.e., observations) and it is interesting to note that we are now using topology as a tool in the study of the fundamentals of measurement and representation of natural systems.

5. f-Continuity

DEFINITION. Let S be a set, T a dynamics on S, and $f \in \mathbb{R}^S$ an observable. Then T is *f-continuous* if T is a continuous dynamics on S with the f-topology (whence the f-topology is coarser than the T-topology).

6. Compatibility

Let T be a dynamics on S and R an equivalence relation on S. Recall (Definition IV.17) that T is *compatible with* R if xRy implies $T(x, t)$ is defined if and only if $T(y, t)$ is defined, and $T(x, t)RT(y, t)$; i.e., for all $t \in \mathbb{R}$ and for all $x \in \operatorname{dom} T_t \subset S$, xRy implies $y \in \operatorname{dom} T_t$ and $T_t(x)RT_t(y)$.

In particular, if $f \in \mathbb{R}^S$, then T is compatible with R_f (or simply T is *compatible with f*) if for all $t \in \mathbb{R}$ and for all $x \in \operatorname{dom} T_t$, $f(x) = f(y)$ implies $y \in \operatorname{dom} T_t$ and $f(T_t(x)) = f(T_t(y))$.

THEOREM 1. Let T be a dynamics on S and $f \in \mathbb{R}^S$. If T is f-continuous, then T is compatible with f.

Proof. If T is f-continuous, then each T_t is continuous on $\operatorname{dom} T_t \subset S$ with the subspace f-topology and $\operatorname{dom} T_t$ is f-open (see Lemma V.2(2)). This means that for all $x \in \operatorname{dom} T_t$ and for all $\varepsilon > 0$ there exists a $\delta > 0$ such that

$$d_f(x, y) = |f(x) - f(y)| < \delta$$

implies $y \in \operatorname{dom} T_t$ and $d_f(T_t(x), T_t(y)) = |f(T_t(x)) - f(T_t(y))| < \varepsilon$. Now let $x \in \operatorname{dom} T_t$ and $y \in S$ be such that $f(x) = f(y)$. Then for every $\varepsilon > 0$, $d_f(x, y) = |f(x) - f(y)| = 0 < \delta(\varepsilon)$, hence $y \in \operatorname{dom} T_t$ and $d_f(T_t(x), T_t(y)) = |f(T_t(x)) - f(T_t(y))| < \varepsilon$; so $|f(T_t(x)) - f(T_t(y))| = 0$, whence $f(T_t(x)) = f(T_t(y))$. Thus T is compatible with f. □

It is clear that the converse of this theorem is not necessarily true. Further, combining Theorems V.2(6) and V.6(1), we have the following theorem.

THEOREM 2. Let T be a dynamics on S and $f \in \mathbb{R}^S$, and let T be f-continuous. Then the quotient dynamics T' induced by T on the metric space $(S/R_f, d_f)$ is a continuous dynamics. □

7. T-Continuity

DEFINITION. Let T be a dynamics on S and $f \in \mathbb{R}^S$. Then f is *T-continuous* if $f: S \to \mathbb{R}$ is a continuous function on S with the T-topology (and $\mathbb{R}$ with the usual topology).

Since the f-topology is the coarsest topology on S rendering f continuous, f is T-continuous if and only if the T-topology is finer than the f-topology.

Trivially, we have

LEMMA. (i) ⇒ (ii) ⇔ (iii), where

(i) T is f-continuous.
(ii) The T-topology is finer than the f-topology.
(iii) f is T-continuous. □

D. Time and Dynamics

8. Time as Real Numbers

A main area of investigation in this study is to determine how we may employ formal dynamical models of natural systems to make spatial and temporal predictions about the systems themselves. So let us examine the modelling relations involving dynamical systems in some detail.

The crucial concept in dynamics is, of course, *time*. The concept of time involves two distinct aspects: *simultaneity* and *temporal succession*. Both of these are intimately involved in our study of dynamical systems. In our definition of dynamics, we have tacitly encoded time as a "set of instants" in the continuum $\mathbb{R}$, and we have made extensive use of the mathematical properties of $\mathbb{R}$. Simultaneity and temporal succession are implied by the total order $\leq$ on $\mathbb{R}$. The fact that $(\mathbb{R}, +)$ is an abelian group is basic to the group property of dynamics [see Definition IV.3(iii)].

In this chapter when we looked at topological dynamics, the topological (metric) properties of $\mathbb{R}$ entered in an essential way. This is precisely where the difference between a "continual dynamical system" $\{T_t : t \in \mathbb{R}\}$ and a "discrete dynamical system" $\{T^n : n \in \mathbb{Z}\}$ lies (cf. Section IV.21). Although $(\mathbb{Z}, \leq, +)$ is also a totally ordered abelian group, time encoded as real numbers has the further bonus of the expression of "temporal approximation" over the encoding as integers.

Note, however, that this distinction between $\mathbb{R}$ and $\mathbb{Z}$ is purely mathematical: that $\mathbb{R}$ is "complete" while $\mathbb{Z}$ has the discrete topology. When it comes down to tackling fundamental problems like "What is time?" and "What is a universal encoding for time?" it may very well turn out that $\mathbb{Z}$ suffices. After all, since we are only dealing with operational definitions (see Section III.1), we cannot in principle (and obviously in practice as well) measure any duration of time that is shorter than the time required for light to pass through a single hydrogen atom. So if we take this (very small, but nonetheless positive) theoretical lower bound as our elementary step of time, then any period of time becomes an integral multiple of this elementary step, and hence it suffices to encode time as $\mathbb{Z}$. But because we need the topological properties of the continuum, we shall not dwell on this point and simply encode time as $\mathbb{R}$.

9. Dynamical Time

Let us revisit the rate equations we considered in Section IV.4 and Example V.2(5). But this time we shall consider two sets of them:

$$dx/dt_1 = f(x), \qquad dy/dt_2 = g(y), \tag{1}$$

where x is in an open subset S_1 of $\mathbb{R}^m$, y is in an open subset S_2 of $\mathbb{R}^n$ (m and n may be different), $f: S_1 \to \mathbb{R}^m$, $g: S_2 \to \mathbb{R}^n$ are continuous, and t_1 and t_2 are "times."

In employing rate equations of the form (1) as encodings of dynamical processes in natural systems, the usual point of view is that, roughly, the differential increments $dx = (dx_1, dx_2, \ldots, dx_m)$ of the state variables are known, and the time differential dt_1 is known, and $f = (f_1, f_2, \ldots, f_m)$ is their ratio. We can take a different point of view and say that the principle of causality (see Section IV.5) together with the differentials dx essentially serve to specify the observables f (see Section IV.2: a "change of state" is an observable). Observables f and differentials dx then together *define* the time differential dt_1. Similarly, for the other system, g and dy together define the time differential dt_2. One expects that the differentials arising from different dynamical systems are in general quite different from one another, and hence there is no reason to believe that the time differentials dt_1 and dt_2 are the same, i.e., every dynamics defines its own (intrinsic) time. So the problem is how these different times are related to a common 'clock time' (extrinsic), which is likely to be different from all of them.

To tackle this problem let us reconsider equations (1) and this time take the point of view that t is simply an arbitrary parameter; for example, we could take it to be arc length along the trajectories. Then we can multiply the equations by some nonvanishing function $a_i: S_i \to \mathbb{R}$ without changing the qualitative properties of the systems. All we do is change the rates at which the trajectories are traversed (relative to some fixed extrinsic time scale). In effect, we are replacing dt_i by $dt = a_i(\cdot)dt_i$. It is through this 'scaling' that we convert different intrinsic times to a 'common time' in terms of which dynamical predictions can be made.

So when we define $\phi \in \mathbf{D}((S_1, D_1), (S_2, D_2))$ via the diagram ($T \in D_1$):

$$\begin{array}{ccc} \operatorname{dom} T & \xrightarrow{\ \phi \times 1_{\mathbb{R}}\ } & \operatorname{dom} \phi T \\ {\scriptstyle T}\downarrow & & \downarrow{\scriptstyle \phi T} \\ S_1 & \xrightarrow[\ \phi\ ]{} & S_2 \end{array}$$

the $1_{\mathbb{R}}: \mathbb{R} \to \mathbb{R}$ that appears is really a map that makes the correspondence $t_1 \mapsto t_2$, via the scaling above to a common time t ($dt = a_1(x)\,dt_1 = a_2(y)\,dt_2$); i.e., the element-chasing diagram

$$\begin{array}{ccc} (x,t) & \longmapsto & (\phi x,t) \\ \big\downarrow & & \big\downarrow \\ T(x,t) & \longmapsto & \phi(T(x,t)) = \phi T(\phi x,t) \end{array}$$

should really read

$$\begin{array}{ccc} (x,t_1) & \longmapsto & (\phi x,t_2) \\ \big\downarrow & & \big\downarrow \\ T(x,t_1) & \longmapsto & \phi(T(x,t_1)) = \phi T(\phi x,t_2) \end{array}$$

But the scaling correspondence $t_1 \leftrightarrow t_2$ is assumed from the very beginning to simplify matters. For further discussions on time scaling, see Richardson and Rosen (1979) and Rosen (1981).

VI. THE CATEGORY OF NATURAL SYSTEMS

> Nature is the realization of the simplest conceivable mathematical ideas.
>
> Albert Einstein

A. The Modelling Relation

1. Introduction

A *natural system*, naturally, is a part of the external world; i.e., it is a member of the entities outside of us, but it can generate under appropriate conditions percepts (sensory impressions) in us. These percepts are then identified with specific properties of the natural system itself.

The adjective *natural* is used to distinguish the system from a "formal" one, which is part of mathematics and hence is a *construct* of our minds. But then one of the primary functions of our minds is to *organize* percepts and establish *relations* among them, which is a *constructive* ability. Consequently, creations of the mind are *imputed* to the external world, and hence we have essentially obtained *formal models of natural systems*. Modelling is in fact a fundamental quality of the mind.

The main objective of our study is to establish relations between the classes of natural and formal systems. The difficulty and challenge in establishing such relations arise from the fact that the two classes are entirely different. A natural system is essentially a bunch of linked qualities coded by the specific percepts that they generate and by the relations that the mind creates to organize them. So a *natural system is never completely known*: We continually learn about such a system as our means of observation and our understanding grow. A formal system, on the other hand, is a creation of our minds, and so we do not learn about a formal system beyond establishing the consequences of our definitions through applications of the usual inferential rules of mathematical logic, and sometimes modifying the initial definitions. The basic task of our study is thus *relating experiments to theory.*

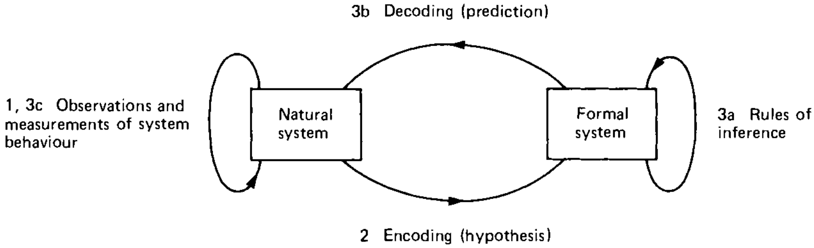
2. Description, Simile, and Metaphor

In *Poetry*, one of the *Six Classics* of Chinese literature, it is stated that there are three major types of figures of speech: *description*, *simile*, and *metaphor*. These three terms, perhaps not too surprisingly, also found their way into the domain of science.

Description in science needs no further explanation; it is the fundamental of experimentation, collection of observable data.

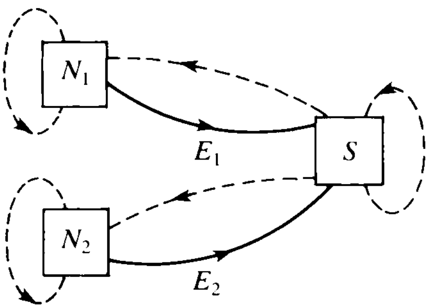
The essential step in our study lies in the exploitation of simile. We are going to force the name of a percept to also be the name of a formal entity, to force the name of a linkage between percepts to also be the name of a relation between mathematical entities, and to force the various temporal relations characteristic of causality in Nature to be synonymous with the dynamical structure of mathematical objects. In short, simile likens these dissimilar things to one another.

Another way to characterize what we are trying to do here is via the three main stages of Bertrand Russell's "scientific process" mentioned in Section I. We seek to *encode* natural systems into formal ones consistently in the sense that the observed phenomena are accounted for, *infer* further consequences from the mathematical structures, and make *predictions* about the natural systems then *verified* when appropriately *decoded*. Diagrammatically, we have



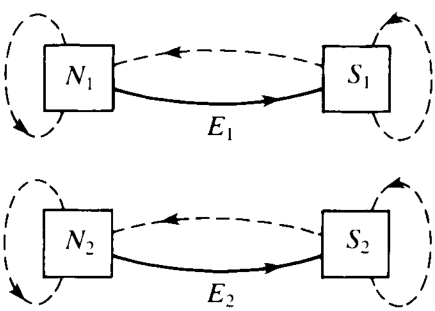
The relation we established between the two preceding systems is the *modelling relation*. We also say that the formal system is a *model* of the natural system, that the latter is a *realization* of the former, and that the two systems are *similes* of each other, via the encoding and decoding rules in question. Note that the idea of simile—the (realization, model) pair—establishes a relation between a natural system and a formal system.

Now suppose we have two natural systems N_1 and N_2 encoded into the same formal system S via encodings E_1 and E_2, respectively:



Then a relation between N_1 and N_2 can be established depending on the "degree of overlap" in S between the "images" $E_1(N_1)$ and $E_2(N_2)$. If $E_1(N_1) = E_2(N_2)$, then the natural systems N_1 and N_2 share a common model, or are *analogues* of each other. *If* $E_1(N_1) \subset E_2(N_2)$, then by restricting to a subsystem N'_2 of N_2 we would have $E_1(N_1) = E_2(N'_2)$, so we say N_2 contains a subsystem analogous to N_1. The more general case of $E_1(N_1) \bigcap E_2(N_2) \neq \varnothing$ can be described as N_1 and N_2 possessing subsystems analogous to each other. Note that the relation of analogy is one among natural systems.

Even more generally, we can consider the situation illustrated in the following (self-explanatory) diagram:



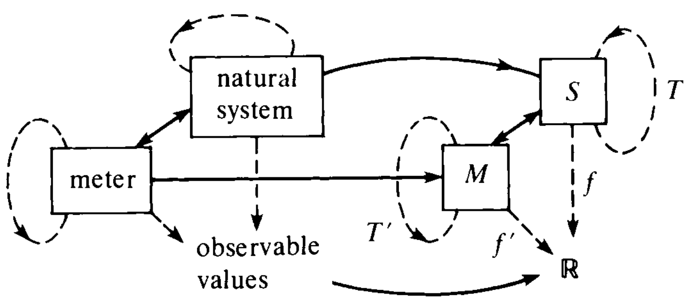
If $E_1(N_1)$ and $E_2(N_2)$ are isomorphic (as mathematical objects in an appropriate category), then via this (not necessarily unique) one-to-one and onto structure-preserving map, we can establish a "dictionary" between the two encodings and hence a relation between N_1 and N_2, called an *extended analogy*. To relax the conditons even further, if $E_1(N_1)$ and $E_2(N_2)$ share some significant general properties in common, a relation can be imputed to N_1 and N_2 themselves; in this case we say that N_1 and N_2 are *metaphors* of each other. Again, the ideas of extended analogy and metaphor are relations among natural systems and can be passed onto subsystems in the same way as in the preceding paragraph.

3. *Observables and Linkages*

Having gone through the preceding discussions, we can now simply say that a natural system is a set of "qualities" on which different definite relations can be imputed. A perceptible quantity of a natural system is obviously what we call an *observable*, and relations among them are *linkages*. The study of natural systems is precisely the specification of its observables, and the characterization of the manner in which they are linked. Thus it becomes clear that the category **S** we looked at in Section III is the appropriate mathematical (formal) tool to be used to study (static models of) natural systems.

Next we have to recognize that (almost by definition) natural systems are dynamic objects and their changes cause a modification in our percepts. Most of the changes in natural systems are of course from their mutual interactions, and in fact the changes in our percepts (these are "observables") can be considered as the result of interactions with other natural systems. So if an interaction between two natural systems causes some change, then the vehicle responsible for the change in one is an observable of the other. This leads us to the discussions of meters and dynamics, and dynamical systems in general, of Section IV. So the category **D** can be used to model the dynamical aspects of natural systems.

We can express these considerations succinctly in a diagram:



In this section we shall be concerned with these models of natural systems. The categories **S** and **D** will be amalgamated into the "category of natural systems," denoted by **N**.

B. The Category N

4. Products of Categories

From two categories **A** and **B** a new category **A** × **B**, called the *product* of **A** and **B**, can be constructed. An object of **A** × **B** is a pair (A, B), where A is an **A**-object and B a **B**-object; a morphism $(A_1, B_1) \to (A_2, B_2)$ of **A** × **B** is a pair (f, g) of **A**-morphism $f: A_1 \to A_2$ and **B**-morphism $g: B_1 \to B_2$; and the composition of two **A** × **B**-morphisms is defined component-wise by

$$(f_2, g_2) \circ (f_1, g_1) = (f_2 \circ f_1, g_2 \circ g_1).$$

Functors $P: \mathbf{A} \times \mathbf{B} \to \mathbf{A}$ and $Q: \mathbf{A} \times \mathbf{B} \to \mathbf{B}$ called *projections* of the product are defined, naturally, by

$$P(A, B) = A, \qquad P(f, g) = f,$$

$$Q(A, B) = B, \qquad Q(f, g) = g.$$

The triple $(\mathbf{A} \times \mathbf{B}, P, Q)$ clearly satisfies the universal property for categorical products and is in fact the **Cat**-product of **A** and **B**.

5. The Category S × D

Objects of **S** × **D** look like $((S, F), (S', D))$ and morphisms in **S** × **D** are $(\phi, \phi'): ((S_1, F_1), (S'_1, D_1)) \to ((S_2, F_2), (S'_2, D_2))$, where $\phi \in \mathbf{S}((S_1, F_1), (S_2, F_2))$ and $\phi' \in \mathbf{D}((S'_1, F_1), (S'_2, F_2))$.

A general **S** × **D**-object (in particular when $S \neq S'$) would be of interest and deserves further study, but for the moment we would concentrate on a subcategory **N** of **S** × **D**, where the **N**-objects are **S** × **D**-objects in which $S = S'$, i.e., of the form $((S, F), (S, D))$, and for convenience it will be denoted simply by (S, F, D).

Let us define the category **N** more explicitly:

6. Objects

An **N**-object is a triple (S, F, D), where (S, F) is an **S**-object and (S, D) a **D**-object.

We shall also consider S as a set on which different topologies can be defined. In particular, we shall consider the topological spaces (S, τ), where τ can be the f-topology for any $f \in F$ and the T-topology for any $T \in D$. Since we assume $0 \in F$ (0 denoting the zero function from S to $\mathbb{R}$) and $I_S \in D$ (the trivial dynamics), S is provided with both the indiscrete and the discrete topology [see V.4 and Example V.2(2)].

7. Morphisms

$$\phi \in \mathbf{N}((S_1, F_1, D_1), (S_2, F_2, D_2))$$
$$\text{if } \phi \in \mathbf{S}((S_1, F_1), (S_2, F_2)) \text{ and } \phi \in \mathbf{D}((S_1, D_1), (S_2, D_2)),$$

i.e., ϕ is a mapping of the sets $S_1 \to S_2$, $F_1 \to F_2$, and $D_1 \to D_2$ such that on (S, F) it satisfies the conditions of III.4 and on (S, D) it satisfies the conditions of IV.9. Note that a general $\mathbf{S} \times \mathbf{D}$-morphism (ϕ, ϕ') on $\mathbf{N}$-objects is not necessarily an $\mathbf{N}$-morphism, because for the latter ϕ and ϕ' have to "agree" on the state (phase) space S. So the inclusion functor $\mathbf{N} \to \mathbf{S} \times \mathbf{D}$ is faithful but not full.

We do not impose any relations for ϕ on (F, D). So, for example, $T \in D_1$ may be f-continuous for some $f \in F_1$, but $\phi T \in D_2$ is not required to be ϕf-continuous; and we can have T compatible with f but ϕT not compatible with ϕf, and so on.

8. Identity and Composition

The identity $\mathbf{N}$-morphism $1_{(S,F,D)}$ is clearly the amalgamation of $1_{(S,F)}$ and $1_{(S,D)}$, i.e., $1_{(S,F,D)}$ sends each $x \mapsto x \in S$, $f \mapsto f \in F$, and $T \mapsto T \in D$.

Composition of $\mathbf{N}$-morphisms is defined component-wise and is clearly associative with identity $1_{(S,F,D)}$.

If $\phi: (S_1, F_1, D_1) \to (S_2, F_2, D_2)$ is an $\mathbf{N}$-isomorphism, then (S_1, F_1) and (S_2, F_2) are $\mathbf{S}$-isomorphic, and (S_1, D_1) and (S_2, D_2) are $\mathbf{D}$-isomorphic (see Sections III.7 and IV.11). Note that even between $\mathbf{N}$-isomorphic systems, the continuity and compatibility properties of the dynamics and observables are not necessarily preserved. This is due to the fact that the observable are less well behaved and that $\sim$-equivalent observables, which are not necessarily topologically equivalent, are $\mathbf{S}$-isomorphic. This apparent shortcoming, contrariwise, turns out to be of great interest; some of these "bifurcation phenomena" will be discussed at the end of this section.

C. Constructions in N

9. Products and Coproducts

Since the phase (state) spaces of the $\mathbf{S}$-product and the $\mathbf{D}$-product of a family are the same (both being the $\mathbf{Ens}$-product), we can construct the $\mathbf{N}$-product simply by "putting the two pieces together." Explicitly, let $\{(S_i, F_i, D_i) : i \in I\}$ be a family of $\mathbf{N}$-objects. Then the $\mathbf{N}$-product of this family is (S, F, D), where (S, F) is the $\mathbf{S}$-product of $\{(S_i, F_i)\}$ (when the $\mathbf{S}$-product exists) and (S, D) is the $\mathbf{D}$-product of $\{(S_i, D_i)\}$ (see Sections III.8 and IV.12).

Because of Theorem V.2(4), we see that continuous dynamics are preserved by $\mathbf{N}$-products. Also, it is easy to see that if the dynamics T_i is compatible with the equivalence relation R_i on S_i, then if the product dynamics $T = (T_i : i \in I)$ exists, it is compatible with the product equivalence relation R on S (where $(x_i : i \in I)R(y_i : i \in I)$ iff for each $i \in I, x_i R_i y_i$). In particular, if T_i is f_i-continuous for each $i \in I$, then $T = (T_i : i \in I) \in D$ is f-continuous ($f = (f_i : i \in I)$) on S, and vice versa.

Similarly, since the phase spaces of the S-coproduct and the D-coproduct are the same (both being the **Ens**-coproduct), the **N**-coproduct of the family $\{(S_i, F_i, D_i) : i \in I\}$ is (S, F, D), where (S, F) is the **S**-coproduct of $\{(S_i, F_i)\}$ and (S, D) is the **D**-coproduct of $\{(S_i, D_i)\}$ (see Sections III.11 and IV.13).

Continuous dynamics are preserved by **N**-coproducts [see Theorem V.2(5)]. Compatibility is also preserved with the coproduct equivalence relation on S defined by $(i, x)R(j, y)$ iff $i = j$ and $xR_i y$ in S_i.

10. Monomorphisms and Subobjects

It is clear that in a product category **A** × **B**, a morphism (ϕ, ϕ') is a mono (an equalizer, split mono,...) if and only if ϕ is mono (an equalizer, split mono,...) in **A** and ϕ' is mono (an equalizer, split mono,...) in **B**. The corresponding statement, with obvious modifications, also holds for any subcategory of **A** × **B**. Further, we can have concepts like "partial monos" in **A** × **B**, where, for example, an **A**-mono in **A** × **B** is an **A** × **B**-morphism (ϕ, ϕ') such that ϕ is an **A**-mono and ϕ' is an arbitrary **B**-morphism.

Thus, an **N**-morphism is mono (an equalizer, split mono,...) if it is such in both **S** and **D**. Now recall the hierarchies for monomorphisms (see Sections III.13 and IV.14).

in **S**: split mono ⇔ equalizer ⇒ mono ⇔ injection,

in **D**: split mono ⇒ equalizer ⇔ mono ⇔ injection.

The hierarchy for **N**-monomorphisms, taking the "intersection" of the two preceding statements, is then

split mono ⇒ equalizer ⇒ mono ⇔ injection (on S, F, and D).

So there are three distinct types of monomorphisms in **N**; hence there are (at least) three possible definitions of an **N**-subobject. We shall choose to call an equalizer-subobject an **N**-subsystem because an equalizer preserves the relevant equivalence class structures on S (see Section III.17).

DEFINITION. Let $\phi \in \mathbf{N}((S_1, F_1, D_1), (S_2, F_2, D_2))$. Then (S_1, F_1, D_1) is an **S**- (respectively, **D**-, **N**-) *subsystem* of (S_2, F_2, D_2) if ϕ is an **S**- (respectively, **D**-, **N**-) equalizer [where ϕ is an **S**-equalizer if $\phi: (S_1, F_1) \to (S_2, F_2)$ is an **S**-equalizer and $\phi: (S_1, D_1) \to (S_2, D_2)$ is an arbitrary **D**-morphism, and so on].

All the different kinds of subsystems in the preceding definition share one decisive feature: the passage from a system to a subsystem places limitations on the interactive capabilities of the system. So an **S**-subsystem is limited in the measurements that can be performed on it (to obtain "observables") and hence is limited in the kinds of dynamics it can impose on other systems; a **D**-subsystem is limited in the kinds of dynamics that can be imposed on it and hence is limited in its capability as a meter; and an **N**-subsystem, which is *both* an **S**-subsystem and a **D**-subsystem, can be limited in both ways.

11. Epimorphisms and Quotient Objects

The hierarchies for epimorphisms are (see Sections III.14 and IV.16)

in **S**: split epi $\Rightarrow$ coequalizer $\Rightarrow$ epi $\Leftrightarrow$ surjection,

in **D**: split epi $\Rightarrow$ coequalizer $\Leftrightarrow$ epi $\Leftrightarrow$ surjection.

Therefore the hierarchy for **N**-epimorphisms is

split epi $\Rightarrow$ coequalizer $\Rightarrow$ epi $\Leftrightarrow$ surjection (on S, F, and D).

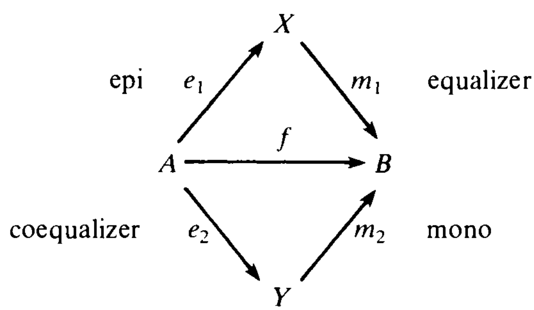
So there are three distinct types of epimorphisms in **N**, and hence many various definitions of **S**-, and **D**-, and **N**-quotient objects. (Incidentally, a *quotient object* is the dual concept to that of a subobject; a subobject is the domain of a monomorphism while a quotient object is the codomain of an epimorphism.)

We emphasize again that even in subobjects and quotient objects, the continuity properties of the dynamics are preserved only under special circumstances [see Lemma V.2(3) and Theorem V.2(6)].

12. Image Factorizations

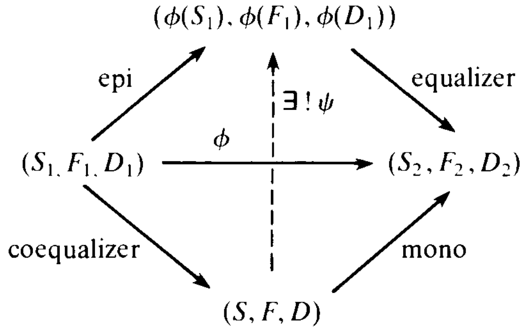
N has both epi-equalizer and coequalizer-mono factorizations. The two factorizations are distinct in general, because of the distinct factorizations in **S** (see Sections III.16 and IV.16).

Let us consider first a general category **C** that has both epi-equalizer and coequalizer-mono factorizations. Let $\mathscr{E}$ = all **C**-epis and $\mathscr{M}$ = all **C**-equalizers. Let $f: A \to B$ in **C**, let $f = m_1 \circ e_1 : A \to X \to B$ be its epi-equalizer factorization, and let $f = m_2 \circ e_2 : A \to Y \to B$ be its coequalizer-mono factorization.



Then e_2 is a coequalizer hence an epi (general hierarchy) so $e_2 \in \mathscr{E}$; and $m_1 \in \mathscr{M}$. Thus by the diagonal fill-in lemma (see Section IV.15) there exists a unique $h: Y \to X$ with $e_1 = h \circ e_2$ and $m_2 = m_1 \circ h$. Further, since $e_1 = h \circ e_2$ is epi, so is h (see Section III.14); and since $m_2 = m_1 \circ h$ is mono, so is h (see Section III.13). In other words, there is a unique **C**-morphism h from the "coequalizer-mono image" Y to the "epi-equalizer image" X and h is both epi and mono (not necessarily an isomorphism, however).

Coming back to our category **N**, we see that for any $\phi:(S_1, F_1, D_1) \to (S_2, F_2, D_2)$, we have the commutative diagram



where ψ is both **N**-epi and **N**-mono but not necessarily an **N**-isomorphism. ψ is, however, a **D**-isomorphism from (S, D) to $(\phi(S_1), \phi(D_1))$ (see Section IV.16).

D. Bifurcations

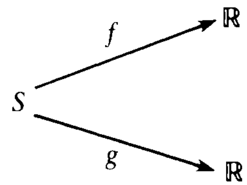
13. Multiple Descriptions

The **N**-object (S, F, D) contains many different mathematical descriptions of the same system. There are many interesting questions about their connections one can ask: How do the descriptions of S obtained from one set of observables $\{f_1, f_2, \ldots, f_n\}$ compare with those from another set $\{g_1, g_2, \ldots, g_m\}$? How can one combine these descriptions to obtain a more comprehensive picture of S? To what extent does a knowledge that two states s_1 and s_2 appear "close" under the pseudometric on S induced by $\{f_1, f_2, \ldots, f_n\}$ imply that these same states appear close with respect to $\{g_1, g_2, \ldots, g_m\}$? How does a dynamics T appear when viewed through an observable f? And, vice versa, how does an observable appear after the passage of a dynamical process?

These questions are intimately related to the notions of *stability* and *bifurcation* (Thom, 1975). However, as we shall now see, questions of this type devolve once again to the notion of linkage (see Section III.9) and thence back to the essence of the modelling relation itself. Let us then take up the question of linkage of observables in more detail.

14. Linkage Reconsidered: Equations of State

Let us begin with the simplest case in which we have a set of states S with two observables $f, g \in \mathbb{R}^S$. We can represent this by a diagram



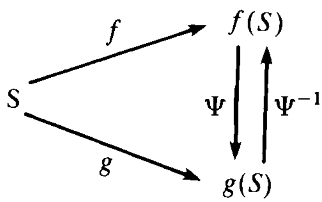
This is yet another type of the diagrams we discussed in Section VI.2. Here a single natural system is encoded into two formal systems, thereby establishing a correspondence between the two latter systems.

The linkage between two observables is, as we indicated in Section III.9, given by how much we learn about $g(s)$ when $f(s)$ is known, and conversely. In other words, linkage between f and g is manifested by the degree of correspondence between the two formal systems ($\mathbb{R}$ and $\mathbb{R}$) encoding S.

The simplest possible situation is when $R_f = R_g$. Then the value $f(s)$ of f at any $s \in S$ completely determines $g(s)$ and vice versa. This means there is actually an (**Ens**-) isomorphism Ψ between $f(s)$ and $g(s)$ [i.e., $(S, \{f\})$ is isomorphic to $(S, \{g\})$ in **S**] and we have for every $s \in S$

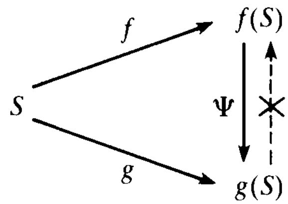
$$\begin{aligned} \Psi(f(s)) &= g(s) \\ f(s) &= \Psi^{-1}(g(s)). \end{aligned} \tag{1}$$

Diagrammatically,



Alternatively, when $R_f = R_g$, we have $R_{fg} = R_f = R_g$ and the image of S/R_{fg} under the embedding $(s)_{fg} \mapsto ((s)_f, (s)_g)$ is simply a "graph" in $S/R_f \times S/R_g$ $(= f(S) \times g(S))$, whose equation is given by (1).

The next situation we can consider is when R_f refines R_g (i.e., when g is totally linked to f). In this case each class in S/R_g is a union of classes in S/R_f, and knowing the value $f(s)$ completely determines $g(s)$, but not vice versa in general. Here $R_{fg} = R_f$ and there is a noninvertible (many-to-one) mapping $\Psi: f(S) \to g(S)$. In other words, the g-encoding may be *reduced* to the f-encoding, but not conversely.



Geometrically, this means that the image of S/R_{fg} in $f(S) \times g(S)$ is a graph satisfying a relation, for all $s \in S$

$$\Phi(f(s), g(s)) = 0, \tag{2}$$

in which $f(s)$ plays the role of an independent variable; i.e., Eq. (2) can be solved for g as a single-valued function of f (e.g., when Φ is differentiable, $\partial\Phi/\partial f \neq 0$) but not conversely (for example, $\partial\Phi/\partial g = 0$).

More generally, if g is partially linked to f, i.e., when the value $f(s)$ of f partially constrains (imposes "selection rules" on) the possible values of $g(s)$, then the image of S/R_{fg} in $f(S) \times g(S)$ still satisfies a (nontrivial) relation of the form (2), except now we can solve for neither f nor g as a single-valued function of the other. Here there is no mapping between encodings in either direction and so neither of them can be regarded as contained in the other.

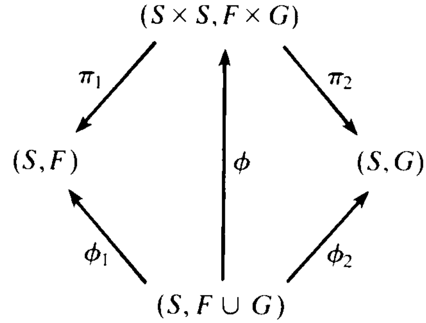
Finally, when f and g are unlinked, the map $S/R_{fg} \to f(S) \times g(S)$ is onto, and there is no meaningful relation of the form (2).

Thus we see that a linkage between two observables f and g can be expressed as a relation of the form (2), which characterizes some subset of $f(S) \times g(S)$. Such a relation is called an *equation of state* for the system. It is a relation between encodings of a natural system, expressing the degree to which the encodings are linked. An equation of state is not an observable but rather represents the encoding of *system laws*; i.e., it is a rule of inference corresponding to a system behaviour.

The preceding discussion can easily be generalized to two families of observables $F = \{f_1, f_2, \ldots, f_n\}$ and $G = \{g_1, g_2, \ldots, g_m\}$ on S. Any linkages between the two families (and in fact within the families) can be expressed by an equation of state of the form

$$\Phi(f_1(s), \ldots, f_n(s), g_1(s), \ldots, g_m(s)) = 0. \tag{3}$$

The character of Φ (e.g., the vanishing of some of its derivatives) will express the manner in which the observables are linked. Categorically, this idea is expressed in the product diagram



and the linkage is reflected in the nature of the S-morphism ϕ.

15. Stable Points

In the preceding section we considered those relationships among observables $f, g, \ldots$ based on the concept of linkage, which depends only on the equivalence classes of the relations R_f, $R_g, \ldots$. We considered the manner in which a "state transition" $s \to s'$ that was undetectable by f (i.e., $f(s) = f(s')$) could be detected by g (i.e., $g(s) \neq g(s')$).

Next, we shall try to answer this question: If s is "close" to s' under (the pseudometric f-topology induced by) f, when will s also be close to s' under g? Thus we are considering the extent to which a state transition $s \to s'$ that is "almost" undetectable by f (i.e., $|f(s) - f(s')|$ is small) is likewise almost undetectable by g (i.e., $|g(s) - g(s')|$ is small).

DEFINITION. Let f, g be observables on S and d_f, d_g be the corresponding induced pseudometrics (see Section V.4). The state $s \in S$ is a *stable point of g with respect to f* if for every $\varepsilon > 0$ there exists a $\delta > 0$ such that for $s' \in S$, $d_f(s, s') = |f(s) - f(s')| < \delta$ implies $d_g(s, s') = |g(s) - g(s')| < \varepsilon$.

The definition is equivalent to each of the following:

(i) the identity map of the set S from (S, d_f) to (S, d_g) is continuous at s;

(ii) the f-open neighbourhood system of s refines the g-open neighbourhood system of s;

(iii) (roughly) every state f-close around s is also g-close around s.

The following is clear:

LEMMA. (a) If $s \in S$ is a stable point of g with respect to f then $(s)_f \subset (s)_g$.

(b) If $s \in S$ is a stable point of g with respect to f and $s' \in (s)_f$, then s' is also a stable point of g with respect to f. □

THEOREM. The set of stable points of g with respect to f is an f-open subset of S. □

16. Bifurcation Set

DEFINITION. The complement of the set of stable points of g with respect to f is the *bifurcation set* of g with respect to f.

Intuitively, near a bifurcation point of g with respect to f, the proximity of two states s and s' of S as viewed by the observable f does not imply their proximity as viewed by g. In other words, the bifurcation set is the set of states at which the g-description does not agree with the f-description in their metrical aspects, i.e., the two descriptions convey essentially different information.

It follows from Theorem VI.15 that the bifurcation set of g with respect to f is an f-closed subset of S. Also, it is clear that the bifurcation set of g with respect to f is empty if and only if the f-topology is finer than the g-topology on S, in which case R_f refines R_g.

17. Equivalence of Observables

In the previous discussion, we can interchange the roles of f and g and obtain the opposite concept of stable and bifurcation points of f with respect to g. These are in general quite different from those obtained from g with respect to f. Thus, given a pair of observables, we obtain two distinct notions of stability and bifurcation, depending on which description is chosen as the reference.

Let us consider the case when f and g are two observables on S such that the bifurcation sets of f with respect to g and of g with respect to f are both empty. Then $1_S:(S, d_f) \to (S, d_g)$ is a homeomorphism and d_f and d_g are equivalent pseudometrics, i.e., f and g induce the same topology on S. Under these circumstances it is appropriate to say that f and g are *topologically equivalent* (as opposed to *algebraically equivalent* $f \sim g$, when $R_f = R_g$). Note that by Lemma VI.15, topological equivalence implies algebraic equivalence; but not conversely. Stated another way, we have the following theorem.

THEOREM. If two observables induce the same topology on the set of states, then they are totally linked to each other. □

18. Bifurcation and Continuity

Let us suppose that the two observables f, g on S are related by an equation of state of the form

$$\Phi(f(s), g(s)) = 0 \tag{1}$$

for all $s \in S$. We shall consider how the concepts of stability and bifurcation are reflected in some (if any) properties of the function $\Phi: f(S) \times g(S) \subset \mathbb{R}^2 \to \mathbb{R}$.

Let $s \in S$ and let U be a neighbourhood of $f(s)$ in $\mathbb{R}$. If $g \circ f^{-1}(U)$ is a small neighbourhood of $g(s)$ whenever U is small, then s is a stable point of g with respect to f. The relation of $g \circ f^{-1}(U)$ to U is clearly another way of expressing the equation of state (1) (around the point $(f(s), g(s))$). It is also clear that s being a stable point of g with respect to f is equivalent to the continuity of the second argument of Φ with respect to the first at $(f(s), g(s))$. Dually, a stable point of f with respect to g occurs at a point of continuity of the first argument of Φ with respect to the second. In other words, bifurcation depends on, and is a consequence of, the discontinuity of the linkage relation between observables, when the properties of distinct encodings become logically independent.

Again, the preceding can be generalized to linkages among more than two observables.

19. Incompatibility

We saw in Section IV.18 that when a dynamics T on a set S is compatible with an equivalence relation R on S, T induces a quotient dynamics on the set of reduced states S/R. In this section we shall investigate what happens when a dynamics T is *not* compatible with an equivalence relation R, in the special case when $R = R_F$ where F is a family of observables on S.

If T is not compatible with R_F, then there are states $s, s' \in S$ for which sR_Fs' but which for some $t \in \mathbb{R}$ T_ts and T_ts' are not R_F-related. That is, T splits equivalence classes of R_F. Putting it another way, the two indistinguishable states s and s' (under F) have now *differentiated* through the action of the dynamics T, and that this differentiation is visible using the observables in F. (The term "differentiation" is used deliberately here to suggest the connection of this with biological differentiation. See Section VII.5.)

From the viewpoint of an observer equipped with meters for observables in F, the states s and s' appeared to be the same. But through the course of the dynamics T the observer detects two different states T_ts and T_ts'. It would appear that the same initial state under the same conditions has given rise to two distinct states, a contradiction to causality. The problem here is of course that one usually assumes that one has a complete set of observables F for the description of S (i.e., $S/R_F = S$, or R_F is the equality relation). The standard way out is to pull in statistics and to observe many copies of $(s)_{R_F}$ under the passage of T. The relative frequencies of the resulting states $(T_ts')_{R_F}$ [where $s' \in (s)_{R_F}$] are then associated with *transition probabilities* from s to T_ts'. In other words, the incompatibility of T with R is usually interpreted in stochastic terms.

One can (rather boldly) make the suggestion that all processes occurring in nature are deterministic, and the apparent stochasticity is simply a consequence of employing an incomplete description as if it were complete. So one does not need statistical tools *if* one has a complete description of the system. But that is a rather big *if* because we are limited in our means of observation, measurement, and understanding, and to obtain a complete description of every natural system is really to find the philosopher's stone! Thus statistics plays a role in science as a matter of necessity. Further discussions of the interplay between causality and chance can be found in Bohm (1957) and in Belinfante (1973) on the theory of "hidden variables."

20. Stability and Bifurcation, Again

We just considered the situation when a change of state $s \to s'$ that is undetectable by an observable f becomes detectable by the observable $g = f \circ T_t$. This of course is again intimately related to the concept of linkage. After all, given an observable f and a dynamics T on S, for each $t \in \mathbb{R}$ $f \circ T_t$ is an observable on S. To say that T is compatible with R_f (i.e., f) simply means that R_f refines each $R_{f \circ T_t}$ or that each $f \circ T_t$ is totally linked to f.

With this in mind, the next natural question to ask is: If s is close to s' under f, when will s also be close to s' under f after the action of a dynamics T? This problem can then clearly be studied by reducing to consider stable and bifurcation points of $f \circ T_t$ with respect to f. And the results from the previous sections can be appropriately modified and used.

Since the compatibility of T with f is equivalent to the commutativity of the following diagram (see the remark of Section IV.18).

$$
\begin{array}{ccc}
\text{dom } T & \xrightarrow{\chi \times 1_{\mathbb{R}}} & \text{dom } T' \\
{\scriptstyle T}\big\downarrow & & \big\downarrow{\scriptstyle T'} \\
S & \xrightarrow[\chi]{} & S/R_f
\end{array}
$$

the study of stability and bifurcation can be formulated in the category $\mathbf{N}$ as the "approximate commutativity" of this diagram.

Finally, there is an interesting possibility that a state $s \in S$ can be a stable point of $f \circ T_t$ with respect to f for all t less than a critical time t_c, and then for $t > t_c$, s becomes a bifurcation point of $f \circ T_t$ with respect to f. In other words, at $t = t_c$ we have a *catastrophe* (in the sense of Thom, 1975). Alternatively we can consider $t \in \mathbb{R}$ as an order parameter and at $t = t_c$ we have a change of scheme from an old structure to a new structure through an instability. This area is a further topic of investigation and we shall not deal with it here. A good set of reference is the Springer Series in Synergetics, especially the introductory Volume 1 (Haken, 1978) and Volume 4 (Güttinger and Eikemeier, eds., 1979), *Structural Stability in Physics.*

VII. BIOLOGICAL IMPLICATIONS

> Any particular or isolated biological phenomenon or group of phenomena admits of necessity an explanation in terms of a mathematical model.
>
> Nicolas Rashevsky

We come now to the final subject of our study. We want to show how the mathematical formalism of categorical system theory we developed may be applied in the analysis of various biological problems.

A. Development and Senescence

1. The Nature of Biology

Biology is the science of life and of the living, and of the multitude of interactions among living systems. The term *life* encompasses an immense set of phenomena; thus biology seems to sit at the rendezvous of all other sciences. The physicist Yang Chen-Ning once said that physics and mathematics are like two leaves sharing the same petiole, that they do not overlap but rather have the same blood in their veins. We can generalize the idea and say that science is like a compound leaf; each leaflet is one branch of science sharing the same rachis, which represents biology, and that biology is a unique science in the sense that this compound leaf is itself a biological object, a living system.

What, then, is a living system and what does it do? Many scientists coming from diverse backgrounds, when engaged in the search for general principles to integrate our understanding of the phenomena of life, have placed major emphasis on the notion of life as viewed from the standpoint of their own specialties. This is a classic example of a reductionist's view of the world. However, although the various theories of living systems differ greatly in their concepts and definitions of basic terms, they seem to share a common goal: to organize the findings (observations!) in some or all of biology into a single conceptual structure.

2. Developmental Biology

The development, the progressive production of the phenotypic characteristics, of organisms is a mysterious process (Berrill and Karp, 1976; Trinkaus, 1969). How can a single cell, the zygote, change itself into an adult containing millions of cells organized in a complex way? The fact that all the information necessary to produce the adult is contained in the very first cell and that the environment simply provides energy and materials increases the difficulty of the problem. The zygote of a given species always becomes an adult of that species and never that of another, whatever its environment. Growth mainly occurs through a continuous process of cell division, but how does differentiation that occurs during growth come about? Since all the cells contain the same genetic instructions, how do they come to differ from one another structurally and functionally and develop into elaborate spatial patterns? How does a developing embryo repair damages that may occur accidentally? Surely precise mechanisms for control and adaptation are involved—but *how*? Perhaps the solution to these complex problems of developmental biology lies not in modelling the phenomena but rather in abstract formal treatments like automata theory (see, for example, Arbib, 1972) and our categorical system theory.

Developmental biology has become an immense field whose boundaries are difficult to define (Thompson, 1942). The focus has been and remains on the embryo—on the gradual emergence of form and structure. The essence of embryonic development is *change*—transition from one state to another (dynamics!). Embryos are a "fleeting state," and development is an expression of the irreversible flow of biological events along the axis of time.

The transformation of an organism or its component cells from one state to another can also be identified in a variety of biological processes not specifically related to embryonic development. Of prime interest among them are aging, regeneration, and malignancy, each of which is characterized by a shift, either gradually or suddenly (catastrophically?), in cellular activities. These processes also fall into the general domain of development biology.

The article by Crick (1977) is an enlightening one about our ignorance—what we do *not* know—of development biology.

3. *Aging*

The outward signs (i.e., observables) of aging are obvious; yet the underlying bases for the deteriorative processes that occur in all living organisms are very poorly understood. In fact, the very definition of the term *aging* is far from clear; it may well be that *aging* is a *primitive* of a logical theory and hence is left as an undefined term. Primitives receive no definitions except those given to them implicitly by their presence in the general axioms of the theory. In our study, the terms *system*, *state*, and *observable* are primitives (see Section III.1). But can we define what aging is in our framework of categorical system theory?

There is no single observable that can be called aging, although many basic observations have been made of age-related changes [i.e., those changes correlated with increasing (chronological) age]. In fact, almost every observable monotone function of chronological age has been used as a measure of biological age. [Note this statement itself assumes that somehow biological age is comparable to chronological age, i.e., time. So just as in dynamical systems (see Section V.9), we are assuming the existence of a "scaling factor" between intrinsic age and extrinsic time.] For example, increasing functions from blood pressure to amount of lipofucsin in cells, and decreasing functions like organ activity, nerve conduction velocity, muscle power, etc., have all been used as age indicators. One of the most important tasks of aging research is to determine which of these changes are primary events that could be directly responsible for survival and which are either secondary manifestations or independent processes not responsible for causing age-related death.

Some investigators believe that aging, whether caused by intrinsic or extrinsic factors, is a general property of all normal cells. Others have proposed that a limited population of cells controls the course of aging throughout the (multicellular) organism. Above all, perhaps it should be emphasized that *aging and growth are inseparable.* Growth as such may be analyzed in terms of cell multiplication, cell enlargement, cell replacement, and other features such as accumulation of extracellular substance. Yet the phenomenon itself, including factors that determine both the rates of growth and the limits of growth, is not understood. Aging is essentially an extension of the growth process into negative values; i.e., the process is seen as increase as long as cell multiplication exceeds cell death and as decrease when cell death exceeds cell multiplication. The process is continuous throughout life, beginning in early development. When the underlying growth-controlling properties become better understood, so will the phenomena of aging and death.

Let us now return to our formal treatment.

4. *A Partial Order in* **N**

DEFINITION. Let (S_1, F_1, D_1) and (S_2, F_2, D_2) be **N**-objects. Then $(S_1, F_1, D_1) \leq (S_2, F_2, D_2)$ if

(i) there is an **N**-mono $\phi: (S_1, F_1, D_1) \to (S_2, F_2, D_2)$;
(ii) S_1, F_1, and D_1 are finite sets.

Note first that if $(S_1, F_1, D_1) \leq (S_2, F_2, D_2)$, then $(S_1, F_1) \leq (S_2, F_2)$ in **S** (see Section III.18). Also, since S_1 is a finite set, for each $f \in F_1$ $f(S_1)$ is a finite subset of $\mathbb{R}$; hence any representative of the class $(f)_\sim$ induces the same f-topology on S_1, i.e., there is a unique $(f)_\sim$-topology on S_1 (see Section V.4).

It is clear that $\leq$ on the finite **N**-objects [i.e., on (S, F, D), where S, F, and D are finite sets] is reflexive and that $\leq$ is transitive. It is also clear, using a similar argument as in Section III.18, that if $(S_1, F_1, D_1) \leq (S_2, F_2, D_2)$ and $(S_2, F_2, D_2) \leq (S_1, F_1, D_1)$ then the **N**-monos involved are in fact **N**-isomorphisms. So $\leq$ is antisymmetric (up to isomorphism, as usual). Therefore $\leq$ is a partial order on (the isomorphism classes of) the finite **N**-objects.

Let $(S_1, F_1, D_1) \leq (S_2, F_2, D_2)$. Since the phase space S_1 is finite, each $T \in D_1$ has dom $T = S_1 \times \mathbb{R}$ (see Corollary 2 in Section IV.6); i.e., for every $x \in S_1$ $a_x[T] = -\infty$ and $b_x[T] = +\infty$. But then from $a_{\phi x}[\phi T] \leq a_x[T] < 0 < b_x[T] \leq b_{\phi x}[\phi T]$, we have $a_{\phi x}[\phi T] = -\infty$ and $b_{\phi x}[\phi T] = +\infty$ as well, i.e., dom $\phi T = S_2 \times \mathbb{R}$. So here we are simply dealing with the usual concept of dynamics. Since dom $T = S_1 \times \mathbb{R} \rightarrowtail S_2 \times \mathbb{R} =$ dom ϕT, T can be considered as the relative dynamics on S_1 induced by the dynamics ϕT on S_2; alternatively, ϕT can be considered as an *extension* of the dynamics T.

5. *Growth and Aging as a Partial Order*

What does this partial order $\leq$ in **N** have to do with our present topic on development and senescence? Since $(S_1, F_1, D_1) \leq (S_2, F_2, D_2)$ in **N** implies in particular that $(S_1, F_1) \leq (S_2, F_2)$ in **S**, we can include the discussions from Section III.18. Let us see what we have.

With an **N**-mono $\phi: (S_1, F_1, D_1) \leq (S_2, F_2, D_2)$, the possibility that ϕ is not onto models *growth*. If $S_2 \sim \phi(S_1)$ is not empty, then more elements (states) have appeared in the second system, an increase in *size*. If $F_2 \sim \phi(F_1)$ is nonempty, then there are more observables in the second system, an increase in *complexity*. If $D_2 \sim \phi(D_1)$ is nonempty, then more dynamics can be imposed on the second system, an increase in *interactive ability*. Collectively, the appearance of these new modes of structure, organization, and behaviour falls into the description of the sometimes puzzling biodynamical phenomenon termed *emergence*.

When $g \in F_2 \sim \phi(F_1)$ and $s, s' \in S_1$ are such that $sR_{F_1}s'$ but $g(\phi s) \neq g(\phi s')$, indistinguishable states in S_1 are now separated in S_2 because of an increase in complexity, an alternate description. When $T \in D_1$ is compatible with R_{F_1} but ϕT is not compatible with some $g \in F_2$ (see Section VI.19), we have the case that indistinguishable states in S_1 become separated in S_2 because of an interaction through an additional dynamics. Both of these cases indicate the presence of *differentiation* going from (S_1, F_1, D_1) to (S_2, F_2, D_2).

Suppose now $s, s' \in S_1$ and $f \in F_1$ are such that $f(s) \neq f(s')$ yet $(\phi f)(\phi s) = (\phi f)(\phi s')$, then distinct states become "the same." Also, when a dynamics is not compatible with an equivalence relation, different equivalence classes may appear to "fuse" together in the course of the dynamical process. These serve as models for biological *integration* when interpreted "positively," and for *decay* (loss of information) when interpreted "negatively."

We mentioned (Section V.9) that the $1_{\mathbb{R}}$ in the map $\phi \times 1_{\mathbb{R}}: \operatorname{dom} T \to \operatorname{dom} \phi T$ is a simplified version of the time scaling in which we have made the identification $t_1 \leftrightarrow t_2$. Suppose instead of $1_{\mathbb{R}}$ we use a monotone increasing function $a: t_1 \mapsto t_2$ from $\mathbb{R}$ to $\mathbb{R}$, then the map $T \mapsto \phi T$ with the commutativity condition $\phi T(\phi x, a(t)) = \phi(T(x, t))$ would have incorporated into it a "speeding up" or "slowing down" of the dynamics depending on whether $\dot{a} > 1$ or $\dot{a} < 1$. This again models growth and decay.

When a change of continuity properties occurs going from (S_1, F_1, D_1) to (S_2, F_2, D_2)—for example, when T is f-continuous ($T \in D_1$, $f \in F_1$) but ϕT is not ϕf-continuous, or when the linkage between ϕf and ϕg is different from that between f and g ($f, g \in F_1$)—it could be interpreted as a change of biological structures and functions. This kind of apparently discontinuous change in biological systems again falls into the area of "emergence." The generation of emergent novelties is highly characteristic of biological systems, and in our formalism of categorical system theory it is a natural consequence of the definition of $\leq$.

With all the preceding in mind, we can now make a formal definition of aging and growth:

DEFINITION. Let (S_1, F_1, D_1) and (S_2, F_2, D_2) be finite natural systems. Then (S_1, F_1, D_1) is *younger* than (S_2, F_2, D_2) (and the latter is *older* than the former) if $(S_1, F_1, D_1) \leq (S_2, F_2, D_2)$; i.e., *aging* is defined as the partial order $\leq$ on the finite natural systems.

6. *The Directionality of Aging*

There is probably no subject that so deeply interests human beings as that of the duration of human life. More than most sciences, gerontology is haunted by primordial myths and fears. This concern is extremely ancient. Somehow people always seem to believe that there are magical ways of reversing, or at least postponing (preferably forever), the process of aging. There are numerous examples in history of searches for these magical ways: Xu Fu was sent by the Chinese emperor in the third century B.C. to look for the Fairy Islands (incidentally, this is supposed to be the origin of Japan); Faust sold his soul to the Devil in exchange for a promise of immortality, knowledge, and power; and the Spanish conquistador Ponce de León searched for the Fountain of Youth in the sixteenth century (but was shot down by native arrows in Florida).

The goals of modern aging research are considerably more modest. We are still at the stage where we are trying to *understand* aging. In fact, the more we know about aging, the more irreversible a process it seems to be. That is why ideas from irreversible thermodynamics and dissipative systems are used to model aging (see, for example, Richardson, 1980). The irreversibility of aging is captured in Definition VII.5. If $(S_1, F_1, D_1) \leq (S_2, F_2, D_2)$ then of course in general we do not have $(S_2, F_2, D_2) \leq (S_1, F_1, D_1)$. Indeed, if the latter holds we would have $(S_1, F_1, D_1) \cong (S_2, F_2, D_2)$ in **N**, in which case the two systems are of the "same age." So the partial order $\leq$ gives a (uni)directionality of aging and growth.

Note, however, that although in general an **N**-mono does not split (see Section VI.10)—this is what we need for $\phi:(S_1, F_1, D_1) \rightarrowtail (S_2, F_2, D_2)$ to be reversible—it is possible that there is a subsystem (S, F, D) of $(\phi(S_1), \phi(F_1), \phi(D_1)) \subset (S_2, F_2, D_2)$ on which ϕ^{-1} exists and is an **N**-mono. In other words, aging results from properties and relations of whole systems, and it does not forbid the possibility that one or more of the component subsystems give opposite contributions. Aging is a *collective* (*cooperative*) *phenomenon* of many processes, some of which may appear to defy aging.

B. The Concept of the Organism

7. *Organism*

When one thinks of collective phenomena in which the discrete constitutive individuals are modified in their behaviour through interactions among one another to fit into the patterns of a larger collective set, and the whole is more than and different from a simple addition of its parts, living organisms would seem to be the ideal example. Yet the concept of the organism has resisted all attempts of definition. Partly it is because organisms are complex, which means that they admit many different kinds of descriptions. The characteristics of any of the many possible descriptions made of an organism provide a possible way of defining the organism. But many of these descriptions look contradictory and hence add to the intrinsic difficulty of the task. The various academic fields that study organisms approach their tasks quite differently. None has a complete picture, and communication among these disciplines is often poor. Gross anatomy concerns structure; general physiology primarily deals with matter–energy processing; and endocrinology, neurophysiology, genetics, and psychology chiefly deal with information processing. On the other hand, since these different descriptions represent different kinds of abstractions from real organisms, each can tell us something that the others cannot. Perhaps the most important problem lies in trying to combine these divergent views of the organism into a more comprehensive picture. (Collective description of a collective phenomenon?)

The conventional (reductionistic) view of the organism is that of a physical-chemical system whose behaviour is to be deduced from its structures according to the laws of physics and chemistry. Life, according to this view, is a potentially self-perpetuating open system of linked organic reactions, catalyzed stepwise and almost isothermally by complex and specific organic catalysts that are themselves produced by the system. Another view of the organism is from that of organic behaviour, in the context of a relational biology, as conceived originally by Nicolas Rashevsky (see Rashevsky, 1960). On the integrative aspects of behaviour, it was Rashevsky's idea that the organisms are recognized as such because we can observe homologies in their behaviours, regardless of the physical structures through which these observations are made. Thus all organisms manifest the same set of basic and ubiquitous biological functions, and through this manifestation organisms could be mapped on one another in such a way as to preserve these basic relations. This idea led to the formulation of Rashevsky's principle of biotopological mappings and Robert Rosen's (categorical) (M, R)-systems (Rosen, 1972). On the adaptive and predictive character of organic behaviour, one is led to the classical (optimal) control theory (Rosen, 1980), and Rosen's theory of anticipatory systems (Rosen, 1985).

The following is a description of the developmental processes of an organism from a categorical standpoint.

DEFINITION. Let **L** consist of

(1) a collection $\{L_i = (S_i, F_i, D_i) : i \in I = [0,1]\}$ of **N**-objects such that for $i \leq j$ in I, $L_i \leq L_j$ in **N**. (So in particular the L_i's are 'finite' **N**-objects.)

(2) for each pair $i, j \in I$ with $i \leq j$, a hom set $\mathbf{L}(L_i, L_j)$ containing a *single* **N**-mono $\phi_{ji}: L_i \to L_j$, such that for $i \leq j \leq k$ in I, $\phi_{kj} \circ \phi_{ji} = \phi_{ki}$

$$L_i \xrightarrow{\phi_{ji}} L_j \xrightarrow{\phi_{kj}} L_k, \qquad L_i \xrightarrow{\phi_{ki}} L_k$$

and if $i > j$ in I then there is no **N**-morphism in **L** from L_i to L_j; i.e., $L(L_i, L_j) = \varnothing$ for $i > j$.

It follows from (2) that for each $i \in I$, $\{\phi_{ii} = 1_{L_i}\} = \mathbf{L}(L_i, L_i)$. Thus **L** is a subcategory of **N** and the set of **L**-objects is totally ordered by $\leq$. In fact, one easily sees that **L** is isomorphic (in **Cat**) to the totally ordered set $(I, \leq)$ considered as a category [see Example II.4(6)].

DEFINITION. An *organism* is a natural system that is (a realization of) **L** for an appropriately chosen family $\{L_i\}$ of **N**-objects and an appropriately chosen collection $\{\phi_{ji}\}$ of **N**-monos satisfying the definition of **L**.

Considering the results we have from the discussions on aging, it is quite reasonable to make the preceding definition of an organism from the standpoint of categorical system theory. The totally ordered set $(I, \leq)$ is an index of *age* and the order $\leq$ on **L** is the process of aging. The instant $i = 0$ can be considered as the moment of *conception* of an organism (when *life* begins) and the instant $i = 1$ is *death*. The developmental processes of the organism are reflected in the systems (S_i, F_i, D_i), the morphisms ϕ_{ji}, and in particular in the evolution of the systems as the index $i \in I$ increases from 0 to 1.

It is an appropriate place here to mention a new subject of study started by Rosen (1985), that of *anticipatory systems*. The basis for this theory is the recognition that most (if not all) of the biological behaviour is of an anticipatory rather than a reactive nature. Let us see if our definition of an organism has incorporated into it this anticipatory character. From the total order on the set of **L**-objects we can conclude that certain things *cannot* happen as the organism develops (i.e., as i increases). For example, it cannot happen that $f(s) = f(s')$ but $(\phi_{ji} f)(\phi_{ji} s) \neq (\phi_{ji} f)(\phi_{ji} s')$, it cannot happen that $T \neq T'$ but $\phi_{ji} T = \phi_{ji} T'$, and so on. So this list of impossible happenings can be considered as a prediction of things to come (or rather, *not* to come). Also, the linkage imposed on the states by the dynamics coupled with causality determines how a present state can be regarded as a model of a future state. There are many qualities like these one can list. So perhaps a formal study of the "category of anticipatory systems" is a perspective for the future. (Note this last statement is itself of an anticipatory character!)

8. Organismic Observables

An observable represents a capacity for interaction manifested by the ability to move a meter. Interactions and indicators are abundant at the level of the organism, especially for human organisms. Let us list a few—the trend is that the more "advanced" the organism, the more observables (where applicable) there are.

Organismic observables (Comorosan, 1976) include measurements of external appearance such as size, posture, colour, and deformity, and external behaviour such as speech and motor functions. They also include such physiological conditions as body temperature, pulse, blood pressure, rate of respiration, brain electrical activity, excretion, sensory functions, and reflexes. There are also very much meter-dependent observables such as appearance under x-ray, fluoroscope, and microscope. (This gives an example of how we continually learn about natural systems as our means of observation and understanding grow.) For highly advanced organisms we can include psychological indicators such as intelligence tests, performance tests, etc. Social organisms can be classified according to social status, roles, and so on. At the present stage of the development of biology, however, the most reliable and precise observables (i.e., those that give us the most complete description of the organism) are probably the biochemical ones. Normal limits and extreme ranges of hundreds of biochemical variables have been determined, methods of measurement specified, instruments perfected, and the meaning of changes in values established. Molecular biology is undoubtedly at the present time the central dogma of biology. Nevertheless, let us not forget what Daniel Bernoulli once said: There is no philosophy which is not founded upon knowledge of the phenomena, but to get any profit from the knowledge it is absolutely necessary to be a mathematician. Mathematical biology shall come of age as a matter of necessity.

C. Organismic Sets and Living Systems

9. Organismic Sets

Organismic sets were built by Nicolas Rashevsky as a representation of biological organisms and societies on a relational basis (Rashevsky, 1972), and a wide range of biological and social phenomena were explained within this framework.

The idea was first started by the observation (!) of the remarkable relational similarities among physics, biology, and sociology. The phenomena are properties of collections of things that are capable of performing certain activities that result in certain products (a cause-and-effect phenomenology!). This led to the suggestion of the existence of a conceptual superstructure of which physics, biology, and sociology are three parallel branches, each partially isomorphic to the other two. (Note the resemblance of the idea of Section VII.1 to this.) This conceptual superstructure is an *organismic set.*

Even before we formally define what an organismic set is, we can see that the setting is perfect for a category theory to be postulated. After all, we are looking at a class of mathematical objects with the same structure. Let us first take a brief digression into considering the idea of *structure* in mathematics in terms of categories.

10. Structured Categories

DEFINITION. Let **C** be a fixed category (usually **Ens**). A *category of* **C**-*objects with structure*, **K**, is given by the following:

(1) **K** assigns to each **C**-object X a class $\mathbf{K}(X)$ of **K**-*structures on* X. A **K**-*structure* is a pair (X,σ) with $\sigma \in \mathbf{K}(X)$.

(2) To each ordered pair of **K**-structures $((X,\sigma),(Y,\tau))$. **K** assigns a subset $\mathbf{K}(\sigma,\tau)$ of $\mathbf{C}(X,Y)$ of **K**-*admissible* **C**-*morphisms from* (X,σ) *to* (Y,τ). For $f \in \mathbf{K}(\sigma,\tau)$ we write $f:(X,\sigma) \to (Y,\tau)$.

They satisfy the following:

(i) *Axiom of composition:* If $f:(X,\sigma) \to (Y,\tau)$ and $g:(Y,\tau) \to (Z,\upsilon)$ then $g \circ f:(X,\sigma) \to (Z,\upsilon)$.

(ii) *Structure is abstract*: If $f: X \to Y$ is a **C**-isomorphism, then for all $\tau \in \mathbf{K}(Y)$ there exists a unique $\sigma \in \mathbf{K}(X)$ such that $f:(X,\sigma) \to (Y,\tau)$ and f^{-1}: $(Y,\tau) \to (X,\sigma)$.

REMARKS. It follows from the definition that $1_X:(X,\sigma) \to (X,\sigma)$ for all $\sigma \in \mathbf{K}(X)$. Thus there is a category, denoted by **K**, with objects all **K**-structures and morphisms all **K**-admissible **C**-morphisms. There is also obviously the *underlying* **C**-*object functor* $F: \mathbf{K} \to \mathbf{C}$ that on objects sends (X,σ) to X and on morphisms sends $f:(X,\sigma) \to (Y,\tau)$ to $f: X \to Y$. When $\mathbf{C} = \mathbf{Ens}$ this functor is of course nothing but the forgetful functor of Example II.8(1).

Because of this $\mathbf{K} \to \mathbf{K}$ association, it is actually sufficient just to study the usual categories and forget about the 'structured categories.' But the latter tends to put things in a better perspective, and so we shall refer to them occasionally.

EXAMPLES. **Gp** and **Top** can be regarded as categories of sets (i.e., **Ens**-objects) with structure, with $\mathbf{Gp}(X)$ = the set of all group structures on the set X and $\mathbf{Top}(X)$ = the set of all topologies on X. The admissible morphisms are chosen appropriately as group homomorphisms and continuous mappings, repectively.

Topological groups can be regarded as a category of groups with structure; here we have $\mathbf{C} = \mathbf{Gp}$ and $\mathbf{K}(X)$ = the set of all topologies on the group X compatible with the group structure, and a **Gp**-morphism is admissible if and only if it is continuous.

To cite a closer example, the category **N** can be considered as a categories of **S**-objects with structure where $\mathbf{N}((S,F))$ is the collection of all sets of dynamics on the phase space S.

11. Organismic Sets

DEFINITION. An *organismic set* is a finite set S such that

(1) Corresponding to each element $e_i \in S$ there is a set S_i^a of *activities* that e_i is capable of performing, and there is a set S_i^p of *products* that e_i can make.

(2) The set of all potential activities $S^a = \bigcup_i S_i^a$ and the set of all products $S^p = \bigcup_i S_i^p$ of the organismic set both have cardinality greater than one.

(3) In a given environment E at a given time t, each e_i only exhibits a proper subset $S_i^a(E,t)$ of S_i^a and makes only a proper subset $S_i^p(E,t)$ of S_i^p. This models *adaptation* and *development* as E and t vary.

(4) S is partitioned into three disjoint subsets S_1, S_2, and S_3 such that $S \sim S_3 = S_1 \cup S_2$ is a "normal" organismic set in itself (i.e., S_3 and its associated S_3^a and S_3^p are the apparently "useless" parts of the organismic set S.), $S \sim S_2$ is an organismic set that can exist but will not develop, and S_1 is the

core of the organismic set so that $S \sim S_1$ cannot exist (i.e., S_1^a and S_1^p are necessary and sufficient for, at least a short range, existence of S).

(5) Taken alone, i.e., removed from S, each $e_i \in S$ has a *survival time* t_i during which they can exist without the availability of S^p. t_i is short compared to the life span of S.

(6) Elements in S^p (i.e., products) act on various e_i's, and so induce a nonempty set of relations S_R with S. It is these relations that make us recognize an organism or a society as such. Members of S_R are in general k-ary relations with $k \geq 2$.

DEFINITION. An organismic set is of *order n* if its elements are organismic sets of order $n - 1$.

We ascribe the order $n = 1$ to the set of genes of cells. Then a multicellular organism is an organismic set of order $n = 2$. A small group of people (e.g., a family) is an organismic set of order $n = 3$. A tribe, as a collection of interacting families, is of the order $n = 4$, and so on. This leads to the idea of *hierarchy* of organismic sets.

The preceding is a very much simplified version of Rashevsky's original formal definition of organismic sets. Attempts to study organismic sets using category theory were made (see Baianu, 1980, for a list references), but ordinary categories did not seem to be sufficient and the theory of supercategories was developed. We shall simply, for our purpose, consider the pieces we took and see if we can fit them into the formalism of our category **N**.

12. Organismic Category

Let **O** be the category of organismic sets; i.e., an **O**-object is an organismic set S. The elements $e_i \in S$ can be considered as "states" of the system S. Members of the sets S^a and S^p can be considered as "observables" on S. In particular the relations in S_R can be interpreted as the observable-induced relations. Although the real-valued requirement of the observables is not met in this case, one can always "digitize" S^a and S^p and impose artificial numbers on them. Of course this digitization has to be done in such a way that there is a minimal loss of information. This situation is perhaps similar to that of numerical taxonomy when we give numerical values to the different taxa. The activities and products in S^a and S^p can also be considered as dynamics induced on the system. Even the survival time t_i of each e_i can be interpreted as the "inherent dynamics" of the element, or as the bounds (a_{e_i}, b_{e_i}). Thus the map $S \mapsto (S, F = S^a \cup S^p, D = S^a \cup S^p \cup \{t_i\})$ is one on the objects of a functor $\mathbf{O} \to \mathbf{N}$. The various other properties of S can be looked upon as further structures on the objects; in other words, **O** can be studied as a category of **N**-objects with structure.

What, then, are the admissible **N**-morphisms ϕ between organismic sets S and S'? Besides the usual requirements of being an **N**-morphism (which implies $\phi: S \to S'$, $S_i^a \to S_i'^a$, $S_i^p \to S_i'^p$, and $\{t_i\} \to \{t_i'\}$), we would like it to preserve the other structures as well. So we want $\phi(S_i^a(E,t)) \subset S_i'^a(E,t)$, $\phi(S_i^p(E,t)) \subset S_i'^p(E,t)$, and $\phi: S_1 \to S_1'$, $S_2 \to S_2'$, $S_3 \to S_3'$, etc. With such a definition, **O**-isomorphic objects would then be abstractly identical organismic sets.

13. Specialized Subsets and Hierarchy

Admittedly, the previous section contains many hand-waving arguments. But the $\mathbf{O} \to \mathbf{N}$ association looks rather promising, and the **O**-morphisms do indeed look like the transformations between abstract diagrams representing biological systems, i.e., the biotopological mappings between graphs of organisms (cf. Rashevsky, 1960). The preceding should be able to acquire mathematical rigor upon further "hard" analysis.

Here we shall just look at two more points. First, the preservation of the "specialized subsets" S_1, S_2, S_3, S^a, and S^p of organismic sets by **O**-morphisms is rather interesting. Recalling that these five sets represent the core, the developer, the nonessential part, the activities, and the products of S, respectively, we can view **O**-morphisms as mappings of different biological functions. Second, when we have a chain of organismic sets $S^1 \to S^2 \to S^3 \to \cdots$ where S^n is of order n, the arrows (ie., **O**-morphisms) in between them are then mappings from one "level" to the next in the hierarchy of organismic sets. These two ideas of "specialized subsets" and "hierarchy" are explored in detail in Miller's (1978) theory of living systems, our next topic.

14. Living Systems

The general living systems theory is presented in Miller (1978) in a conceptual setting concerned mainly with "concrete" systems that exist in space-time. It is not a mathematical treatment but it does seem to provide an (exhaustive) catalogue of what we know about biological and social systems (up to 1978).

Miller classifies complex systems that carry out living processes into seven hierarchical levels—cell, organ, organism, group, organization, society, and supranational system (cf. the orders $n = 1$–7 of organismic sets). The central thesis is that living systems at all of these levels are open systems composed of subsystems that process inputs, throughputs, and outputs. There are 19 critical subsystems essential for life, some of which process matter or energy (named suggestively ingestor, distributor, converter, producer, matter–energy storage, extruder, motor, and supporter), some of which process information (namely, input transducer, internal transducer, channel and net, decoder, associator, memory, decoder, encoder, and output transducer), and some of which process all three (reproducer and boundary). Together they make up a living system. For a list of what all these critical subsystems are at each of the seven levels of living systems, refer to Table 13-1 of Miller's book.

What we want to do here is merely to point out how much one can achieve from two simple postulates (that of critical subsystems—specialized subsets—and that of hierarchy) and a large collection of observed data. We can see how close Miller's theory of living systems is to Rashevsky's theory of organismic sets, and hence the former can somehow also be put into the framework of our categorical system theory. Here we probably want to map each of the critical subsystems to corresponding ones between systems, and "hierarchical morphisms" can also be considered.

Finally, let us conclude these few sections on the "conceptual superstructure" of physical, biological, and social systems by saying that the relations among the components of these systems are not put in there by the imagination of the observer, as shepherds idly trace out a scorpion in the stars. These relations are inherent in the systems and are empirically discovered by the scientists. They are there, patterning the coacting reality, regardless of whether they are observed or not. We must pass beyond raw empiricism, beyond the provincialism of the disciplines, and learn from the cooperation of experiments and theories—how theories are created to stimulate further experiments and how experiments are designed to evaluate theories. The better we can formulate a problem, the more we will know about where and what to look for, and the closer we will get to obtaining a complete set of observables.

D. Description Spaces

15. The Response Tensor and Description Space

The concept of the response tensor was originally presented (Richardson, 1980) in the context of irreversible thermodynamics. In that paper a metric algebra based upon the dissipation function associated with a system was introduced and a measure of the aging of the system was derived from this algebra. The mathematical formalism underlying the concepts of the response tensor and the space it spans called the description space was then presented as a phenomenological calculus for complex systems in Richardson, Louie, and Swaminathan (1982).

In these few sections we shall show how these ideas are connected to those of categorical system theory. Let us first have a quick review of the basic definitions and postulates of the phenomenological calculus.

DEFINITION. Let V be the vector space $\mathbb{R}^n$ equipped with the standard inner (dot) product. Let $\mathbf{F}$ be a (contravariant) vector in V and $\mathbf{a}$ be a (covariant) vector in the dual space V^* (the space of all linear functionals on V). Then the *dyad* $\mathbf{R} = \mathbf{aF}$ is a bilinear function from $V^* \times V$ to $\mathbb{R}$ (i.e., $\mathbf{R}$ is in $T_1^1(V)$, the tensor space of type $(1,1)$ over V) defined by $\mathbf{R}(\mathbf{a}', \mathbf{F}') = (\mathbf{a} \cdot \mathbf{a}')(\mathbf{F} \cdot \mathbf{F}')$. A finite linear combination of dyads is called a *dyadic*, whose action on $V^* \times V$ is defined the natural way (linearly on the images); so a dyadic is also a bilinear function from $V^* \times V$ to $\mathbb{R}$, i.e., in $T_1^1(V)$.

It follows from dimensional arguments that the collection of all dyadics is actually the whole of $T_1^1(V)$. In other words, each tensor of type $(1,1)$ has a representation as a dyadic.

DEFINITION. Let $\mathbf{R} = \Sigma_i \mathbf{a}_i \mathbf{F}_i$ and $\mathbf{S} = \Sigma_j \mathbf{b}^j \mathbf{G}_j$ be two dyadics. Their *double dot product* is the real number $\mathbf{R} : \mathbf{S} = \Sigma_i \Sigma_j (\mathbf{a}^i \cdot \mathbf{b}^j).(\mathbf{F}_i \cdot \mathbf{G}_j)$.

One can check that the double dot product is independent of the representation of the dyadics and is in particular independent of the basis chosen for V. Also, it is easy to see that : is a positive definite bilinear form. This gives the following theorem.

THEOREM. $(T_1^1(V), :)$ is an inner product space. □

The phenomenological calculus arising from the mathematical structure of $(T_1^1(V), :)$ is based upon three postulates:

POSTULATE 1. A given system is characterized by a set of vectors $\{\mathbf{a}^i : i = 1, 2, \ldots, m\}$ in (the dual space of) $\mathbb{R}^n$, and this set depends on the physical constitution of the system. As far as describing the dynamic response of the system to the imposition of a set of forces (or more generally, *causes*) $\{\mathbf{F}_i : i = 1, 2, \ldots, m\}$ in $\mathbb{R}^n$, they form a complete set of constitutive parameters. The index i denotes subsystems of the system.

POSTULATE 2. The system dynamics are characterized phenomenologically by the dyadic response tensor $\mathbf{R} = \Sigma_i \mathbf{a}^i \mathbf{F}_i$.

POSTULATE 3. The space spanned by $\mathbf{R}$, i.e., for fixed covariant vectors $\mathbf{a}^1, \ldots, \mathbf{a}^m$ in $V^* = \mathbb{R}^n$, the set $[\mathbf{a}^1, \ldots, \mathbf{a}^m] = \{\mathbf{R} = \Sigma_i \mathbf{a}^i \mathbf{F}_i : \mathbf{F}_i \in V = \mathbb{R}^n\}$, is called the description space, and $\mathbf{R}$ is invariant under coordinate transformations.

Thus the description space $[\mathbf{a}^1, \ldots, \mathbf{a}^m]$ is a linear subspace of $T_1^1(V)$ and hence $([\mathbf{a}^1, \ldots, \mathbf{a}^m], :)$ is an inner product space. The condition $|\mathbf{R}| = (\mathbf{R} : \mathbf{R})^{1/2} \geq 0$ for a response tensor is interpreted as a principle of directionality for dissipation, aging, and in fact all cause-and-effect phenomenologies.

Instead of the Euclidean inner product space $V = (\mathbb{R}^n, \cdot)$, we could have used a general Hilbert space H and all of the mathematics would still go through. In the special case of $H = L^2$ we would have incorporated into the setting the "time-dependence" of the causes $\mathbf{F}(t) \in L^2$ and the "cause-dependence" of the constitutive parameters $\mathbf{a}(\mathbf{F}) \in (L^2)^* = L^2$. This "Hilbert description space" turns out to be a very interesting mathematical object, the study of which is certainly a worthy project.

16. The Description Space as a Natural System

Let S be an open subset of the Euclidean space $V = \mathbb{R}^n$. A linear functional $\mathbf{a}$: $\mathbb{R}^n \to \mathbb{R}$ ($\mathbf{a} \in V^*$) can be considered as a real-valued function, i.e., an observable, on the state space S. So the set of constitutive parameters (or "coordinate vectors") $\{\mathbf{a}^i\}$ of the description space $[\mathbf{a}^1, \ldots, \mathbf{a}^m]$ can be considered as the set of observables on S; i.e., $F = \{\mathbf{a}^i\}$.

Recall [see Section IV.4 and Example V.2(5)] that a vector field $\mathbf{F}: S \to \mathbb{R}^n$ defining an autonomous differential equation

$$dx/dt = \mathbf{F}(x) \tag{1}$$

gives rise to a C^1 (continuously differentiable)-dynamics T where $T(x, t) = y_x(t)$ is the unique solution to (1) satisfying $T(x, 0) = y_x(0) = x$. It is interesting to note that the converse also holds, namely, given a C^1-dynamics T: dom $T \subset S \times \mathbb{R} \to S$ (where S and $\mathbb{R}$ have the usual topology), there is associated with it a vector filed and hence an autonomous differential equation. Define $\mathbf{F}: S \to \mathbb{R}^n$ by

$$\mathbf{F}(x) = \frac{d}{dt} T(x, t)|_{t=0} \tag{2}$$

then for $x \in S$, $\mathbf{F}(x)$ is a vector in $\mathbb{R}^n$, which we can think of as the tangent vector to the T-trajectory $y_x(a_x, b_x)$ at $t = 0$. And it is clear that y_x is the unique solution to the autonomous differential Eq. (1) satisfying the initial condition $y_x(0) = x$. Thus this establishes a correspondence between vector fields $\mathbf{F}$ and C^1-dynamics T on S. The collection of m-tuples of causes (or "components") $\{\mathbf{F}_i\}$ defining response tensors $\mathbf{R} = \Sigma_i \mathbf{a}^i \mathbf{F}_i$ can then be interpreted as C^1-dynamics on the phase space S through this correspondence. This family of C^1-dynamics on S is then considered as the set D for the natural system (S, F, D).

So we have established that a description space over $V = \mathbb{R}^n$ is in fact a special kind of natural systems fully equipped with its sets of observables and dynamics. This embedding $[\mathbf{a}^1, \ldots, \mathbf{a}^m] \mapsto (S, F, D)$ is quite remarkable in that even the physical interpretation of the different corresponding entities coincides. The set of constitutive parameters of a description space and the set of observables of a natural system are both indicators of the *complexity* of the system, and the former set is actually what we observe on a physical system (by Postulate 1 of the previous section). The connection $F \leftrightarrow T$ between V and D is even more transparent: a force field is the time derivative of a dynamics [in the sense of Eq. (2)] in classical physics.

17. The Category of Description Spaces

Now that we have found out that description spaces are representable as **N**-objects, the next natural question to ask is, "What are the **N**-morphisms between description spaces that would preserve all the relevant structures?" We can consider description spaces as **N**-objects with structure and look for admissible **N**-morphisms (see Definition VII.10). But because of the $\mathbf{K} \to \mathbf{K}$ association and the underlying object functor $\mathbf{K} \to \mathbf{C}$ mentioned in Section VII.10, we can simply consider the collection of all description spaces as a usual category equipped with a functor to **N**. This functor takes a description space to its associated **N**-object, and an **N**-morphism between description spaces to one between **N**-objects.

Let us call the category of description spaces **R**, i.e., an **R**-object is a description space $[\mathbf{a}^1, \ldots, \mathbf{a}^m]$ (over the inner product space $V = \mathbb{R}^n$). Since a description space has a "linear" structure, we would like an **N**-morphism between two description spaces $\phi: [\mathbf{a}^1, \ldots, \mathbf{a}^m] \to [\mathbf{b}^1, \ldots, \mathbf{b}^l]$ to be "linear" as well:

$$\phi\left(\sum_i \mathbf{a}^i \mathbf{F}_i\right) = \sum_i \phi(\mathbf{a}^i)\mathbf{F}_i \tag{1}$$

and so it is clear that ϕ is determined upon further restriction of the m images $\phi(\mathbf{a}^1), \ldots, \phi(\mathbf{a}^m) \in \{\mathbf{b}^1, \ldots, \mathbf{b}^l\}$ to satisfy (1). The linearity condition (1) is a restriction of ϕ on the observables. This is analogous to the situation in linear algebra where a linear transformation is uniquely determined by its action on a basis of the domain vector space. So the set of coordinate vectors $\{\mathbf{a}^i\}$ does indeed behave like a basis although it is not a basis in the vector space sense. By virtue of ϕ being an **N**-morphism and the linearity condition (1), if, for example, $\mathbf{a}^3 = c_1\mathbf{a}^1 + c_2\mathbf{a}^2$ for $c_1, c_2 \in \mathbb{R}$, then we must have $\phi(\mathbf{a}^3) = c_1\phi(\mathbf{a}^1) + c_2\phi(\mathbf{a}^2)$. Note that the relation between **R** and **N** is astonishingly similar to that between **Vect** and **Ens**, where **Vect** is the category of vector spaces and linear transformations over a fixed field.

18. Special **R**-*Morphisms*

It is shown in Richardson, Louie, and Swaminathan (1982) that if the coordinate vectors $\{\mathbf{a}^i\}$ span a subspace of V^* of dimension $k(\leq n)$, then the description space $[\mathbf{a}^1, \ldots, \mathbf{a}^m]$ is of dimension kn (over $\mathbb{R}$). Now suppose the coordinate vectors $\{\mathbf{b}^j\}$ of a second description space $[\mathbf{b}^1,\ldots,\mathbf{b}^l]$ also span a subspace of V^* of dimension k, then $[\mathbf{b}^1,\ldots,\mathbf{b}^l]$ is again of dimension kn and so we would expect somehow that the two description spaces are "isomorphic."

Now what is an **R**-isomorphism? Clearly, it has to be an **N**-isomorphism in the first place. So far we have neglected the double dot product on the description spaces. Recalling that a linear transformation between two inner product spaces of the same (finite) dimension is an isomorphism if and only if it preserves inner products, we shall say that an **R**-morphism ϕ *preserves double dot products* if for all response tensors **R** and **S**,

$$\mathbf{R}:\mathbf{S} = \phi(\mathbf{R}):\phi(\mathbf{S}) \tag{2}$$

[Note the two double dot products appearing on the two sides of Eq. (2) are on different description spaces.] Then we shall say that the two description spaces $[\mathbf{a}^1,\ldots,\mathbf{a}^m]$ and $[\mathbf{b}^1,\ldots,\mathbf{b}^l]$ (of the same dimension kn) are **R**-isomorphic if there is an **R**-morphism that is an **N**-isomorphism and preserves double dot products between the two spaces. So under this definition, **R**-isomorphic description spaces are abstractly the same with respect to all of their mathematical structures.

Next, suppose $\phi \in \mathbf{R}([\mathbf{a}^1,\ldots,\mathbf{a}^m], [\mathbf{b}^1,\ldots,\mathbf{b}^l])$ is such that there exists an $\varepsilon > 0$ and for every $\mathbf{R} \in [\mathbf{a}^1,\ldots,\mathbf{a}^m]$, $|\mathbf{R} - \phi(\mathbf{R})| < \varepsilon$. This condition can be roughly stated as $|\mathbf{R}(\mathbf{a}) - \mathbf{R}(\mathbf{b})| < \varepsilon$ in which the notation is self-explanatory. This leads us to the idea of the "distance" between response tensors from different description spaces. It is intuitively clear that the closer two description spaces are to being "identical," the smaller the norm $|\mathbf{R}(\mathbf{a}) - \mathbf{R}(\mathbf{b})|$ will be. And conversely the smaller the norm, the more **R**-isomorphic the two spaces are. So while the condition $|\mathbf{R}| \geq 0$ describes the dissipation (i.e., aging) *within* a system, the condition $|\mathbf{R}(\mathbf{a}) - \mathbf{R}(\mathbf{b})| \geq 0$ allows one to compare the extent of aging *between* two systems. The former depends only on the constitution (i.e., structure) of a system itself, and the latter depends on the morphisms (i.e., on how close they can get to being identities) between systems.

There is, moreover, an alternate description of the intersystem comparision of aging. As usual, **R**-monomorphisms give rise to a partial order on the **R**-objects, where an **R**-monomorphism is some natural analogue of an injective linear transformation and an **N**-monomorphism. Thus $|\mathbf{R}(\mathbf{a}) - \mathbf{R}(\mathbf{b})|$ gives an indication of how close two systems are in age while $\mathbf{R}(\mathbf{a}) \leq \mathbf{R}(\mathbf{b})$ gives an ordering, a directionality to aging. This bears a remarkable resemblance to the aspects of *simultaneity* and *temporal succession* in the concept of time we discussed in Section V.8. Perhaps this is not too surprising. After all, although aging and time are distinct concepts, they do share a lot of characteristics in common. In particular, they are both clocks—aging is an intrinsic clock and time is an extrinsic clock—for natural systems.

19. Recent Developments

The epistemological exploration of the phenomenological calculus has been continued since the 1982 paper. The sequence of additional publications so far consists of Louie, Richardson and Swaminathan (1982), Louie and Richardson (1985), Richardson and Louie (1983), Richardson and Louie (1985), and Louie and Richardson (1985). The phenomenological calculus turns out to be a general algorithm for the synthesis of mathematical representations of complex, highly interacting systems, and the metric structure inherent in the algorithm provides relationship-connecting representations.

Topics discussed under the framework of the phenomenological calculus include recognition processes, duality and invariance, projective representations, irreversible thermodynamics, quantum mechanics, relativity, and information. The list is surprisingly diverse.

VIII. CONCLUSION

And the end of all our exploring
Will be to arrive where we started
And know the place for the first time.

T. S. Eliot, *Little Gidding*

Our investigation of natural systems started with the reciprocity between observables and dynamics, which reflects the duality of the structural and functional aspects of a system. The formal study is based on the encoding of these static and dynamic realities into mathematical objects in the appropriate categories, and above all based on the two fundamental propositions of Section III.1.

Although Bertrand Russell's idea on the *scientific process* provides a framework on which to discuss the modelling relation (see Section VI.2), which is crucial to our treatment, the term "scientific process" is somewhat inadequate. It is true that we do learn from observation and collection of experimental data, that we do try to theorize to explain these data, and that we do correlate between experiments and theories in order to better both. But this is *modelling*, which only forms a *part* of the scientific process. Science is more than that. Especially since Albert Einstein, the idea of *creative imagination* enters science in an essential way. Innovations in science are consequences more of the genius of the human brain than of "trial-and-error" modelling. So if the modelling relation diagrams in Section VI.2 were to represent the scientific process, on them we should perhaps superimpose pictures of the human brain. As the nineteenth-century scientist Baron Justus von Liebig said, "Every property of an object may give, under appropriate circumstances, a key to a locked door; but theory is the master key that opens all doors." And abstract theories not directly constructed from models of the observed phenomena are probably the best keys.

Having said that natural systems are parts of the external world hence real while formal systems are creations of our minds hence abstract (see Section VI.1), we should recognize that the distinctions between reality and abstraction are not that clear. The real can simply be considered as a consequence of the abstract: "So the living and the dead, things animate and inanimate, we dwellers in the world and this world wherein we dwell... are bound alike by physical and mathematical law" (D'Arcy Wentworth Thompson). On the other hand, the abstract can simply be considered as a piece of the real: "I believe that mathematical reality lies outside us, and that our function is to discover or observe it, and that the theorems which we prove and which we describe grandiloquently as our creations are simply our notes of our observations" (Godfrey Harold Hardy). These considerations reinforce one of the points we try to make in this study, namely, that the abstract properties of measurement pertain as much to pure mathematics as they do to natural systems, and conversely, that pure mathematics is in a sense a result of our measurements.

In Section VII we saw how categorical system theory can be applied to analyze biological problems. The exploration of each area turned out to be the study of a structured category on the category of natural systems. We examined these categories individually but we did not really consider the relations among them. Recall that there are, roughly speaking, three levels in the hierarchy within category theory: namely, categories, functors, and natural transformations. We certainly did use categories a lot, and functors were also employed occasionally, but we did not really talk about natural transformations other than just defining them (see the Definition in Section II.11). Since functors are used to "relate" categories and the functors between two categories are the objects of the "functor category" (see Section II.12) in which the morphisms are natural transformations, and since natural transformations are intimately related to the concept of *similarity* [Rosen (1978)], it is certainly worthwhile to look into these further. Also , the categories **Top** and **Vect** appeared in various places in this study, and it would be fruitful to look for the connections between these and our "system categories" in more detail.

In sum, this work represents a formal extension of some of the ideas suggested in Rosen (1978). I have expressed the fundamentals of measurement and representation of natural systems in the setting of the abstract mathematical theory of categories. I believe I have achieved in this work, with perhaps the exception of the last section, a level of mathematical rigour that is recognizable as such. In this Section VII, although the mathematics is not as stiff, I have managed to weave together several theories of natural (alias biological, aging, dissipative, organismic, living, complex...) systems from the standpoint of our categorical system theory. So this categorical extension of Rosen (1978) does indeed look promising. This extension is, however, not unique, and there are many areas that can be further explored. These will be some of my projects for the years to come.

Once I came across a passage written by A. Lawrence Lowell; since then it has become the directive of my scientific life:

"You will seek not a near, but a distant, objective, and you will not be satisfied with what you may have done. All that you may achieve or discover you will regard as a fragment of a larger pattern, which from his separate approach every true scholar is striving to descry."

ACKNOWLEDGMENTS

My thanks are due to Professor Robert Rosen, mathematical biologist, my supervisor, teacher, and friend, who taught me, among other things, the importance of pluralism.

I am grateful to Professor I. W. Richardson for his generous support, numerous inspiring discussions, and valuable suggestions on how to improve the manuscript.

I would like to extend my appreciation to my colleagues in the Biomathematics Annex—our Red House—for their friendliness and many helpful comments. Special thanks go to E. A. D. Foster for a critical reading of the first draft of this thesis.

My thanks also go to Professor W. R. S. Sutherland, my advisor from the Department of Mathematics, for his guidance.

I acknowledge the Killam Trust of Dalhousie University for support in the form of an Izaak Walton Killam Memorial Scholarship.

Above all, it is a pleasure to express my deepest gratitude to my wife, Margaret, without whose encouragement and understanding nothing could have been done.

REFERENCES

Arbib, M. A. (1972). Automata theory in the context of theoretical embryology. "Foundations of Mathematical Biology," Vol. 2, Chap. 3, 141–215. Academic Press, New York.

Baianu, I. C. (1980). Natural transformations of organismic structures. *Bull. Math. Biol.* **42**, 431–446.

Belinfante, F. J. (1973). "A Survey of Hidden-Variable Theories." Pergamon Press, New York.

Berrill, N. J. and Karp, G. (1976). "Development." McGraw-Hill, New York.

Bertalanffy, L. V. (1968). "General System Theory." Braziller, New York.

Bohm, D. (1957). "Causality and Chance in Modern Physics." Van Nostrand-Rheinhold, Princeton, New Jersey.

Comorosan, S. (1976). Biological observables. *Prog. Theor. Biol.* **4**, 161–204.

Crick, F. H. C. (1977). Developmental biology "The Encyclopaedia of Ignorance," 299–303. Pergamon Press, London.

Eddington, A. S. (1939) "The Philosophy of Physical Science." Cambridge Univ. Press, London.

Eddington, A. S. (1949) "Fundamental Theory." Cambridge Univ. Press, London and New York.

Eilenberg, S. and Mac Lane, S. (1945) General theory of natural equivalences. *Trans. Am. Math. Soc.* **58**, 231–294.

Güttinger, W. and Eikemeier, H. ed. (1979). "Structural Stability in Physics." Springer-Verlag, Berlin and New York.

Hájek, O. (1968). "Dynamical Systems in the Plane." Academic Press, London.
Haken, H. (1978). "Synergetics: An Introduction," 2nd ed. Springer-Verlag, Berlin and New York.
Jeans, J. (1945). "The Astronomical Horizon." Oxford Univ. Press, London and New York.
Louie, A. H., and Richardson, I. W. (1983). *Math. Modelling* **4**, 555–565.
Louie, A. H., and Richardson, I. W. (1985). *Math. Modelling* **6**, (in press).
Louie, A. H., Richardson, I. W., and Swaminathan, S. (1982). *J. Theor. Biol.* **94**, 77–93.
Mac Lane, S. (1971). "Categories for the Working Mathematician." Springer-Verlag, Berlin.
Manes, E. G. (1976). "Algebraic Theories." Springer-Verlag, Berlin and New York.
Miller, J. G. (1978). "Living Systems." McGraw-Hill, New York.
Rashevsky, N. (1960). "Mathematical Biophysics." 3rd rev. ed. Dover, New York.
Rashevsky, N. (1972) "Organismic Sets." Mathematical Biology, Inc., Holland, Michigan.
Richardson, I. W. (1980). The metrical structure of aging (dissipative) systems. *J. Theor. Biol.* **85**, 745–756.
Richardson, I. W., and Louie, A. H. (1983). *J. Theor. Biol.* **102**, 199–223.
Richardson, I. W., and Louie, A. H. (1985). *Math. Modelling* **6**, (in press).
Richardson, I. W., and Rosen, R. (1979). Aging and the metrics of time. *J. Theor. Biol.* **79**, 415–423.
Richardson, I. W., Louie, A. H. and Swaminathan, S. (1982). A phenomenological calculus for complex systems. *J. Theor. Biol.* **94**, 61–76.
Rosen, R. (1958). The representation of biological systems from the standpoint of the theory of categories. *Bull. Math. Biophys.* **20**, 317–342.
Rosen, R. (1971). "Dynamical System Theory in Biology." Wiley, New York.
Rosen, R. (1972). Some relational cell models: the metabolic-repair system. "Foundations of Mathematical Biology," Vol. 2, Chap. 4, 217–253. Academic Press, New York.
Rosen, R. (1976). Structural stability, alternate descriptions and information. *J. Theor. Biol.* **63**, 19–31.
Rosen, R. (1978). "Fundamentals of Measurement and Representation of Natural Systems," Elsevier, New York.
Rosen, R. (1980). On control and optimal control in biodynamic systems. *Bull. Math. Biol.* **42**, 889–897.
Rosen, R. (1985) "Anticipatory Systems." Elsevier, New York. In Press.
Russell, B. (1931) "The Scientific Outlook." Norton, New York.
Thom, R. (1975). "Structural Stability and Morphogenesis." Benjamin, Reading, Massachusetts.
Thompson, D. W. (1942). "On Growth and Form," 2nd ed. Cambridge Univ. Press, London.
Trinkaus, J. P. (1969). "Cells into Organs." Prentice-Hall, Englewood Cliffs, New Jersey.

Index

D

E

F

K

L

M

N

O

P