

On Critical Doctrine of Method in Brain-theory

I. Introduction: The Essential Subjectivity of Comparison

In beginning the development of a theory of brain-object using a top-down approach to mind-brain science, the first problem to be faced is one of developing research methodology. Between the starting point in mental physics and the ending point (a theory of *soma*), it can be reasonably anticipated that many intermediate steps will be required. This is regardless of whether the approach methodology employed is scientific reduction, model order reduction, or combination of the two. As Kant says in *Critique of Pure Reason*, proper Critical doctrine of method requires three things:

1. **discipline** – the compulsion through which the constant propensity to stray from fixed rules is curtailed and finally extirpated (B: 737);
2. **canon** – the embodiment of *a priori* fundamental principles of the correct use of a sure overall faculty of knowledge (B: 824); and
3. **architectonic** – the art of systems (B: 860).

Kant also says **history** is part of the doctrine of method. He did not provide a doctrinal treatment of this, saying it was a place left open in his system and requiring a later filling in. In *Critique of Pure Reason* what he did provide was a very brief summary recap of steps and approaches in the history of philosophy. As discipline, canon, and architectonic align with the transcendental topics of Quantity, Quality, and Relation, respectively, Kant's history requirement belongs to the topic of Modality. It adds nothing to object methodology and rather speaks in regard to *judgment* of methodology, just as Modality in judgment is judgment of the judgment rather than of the object. Modality in transcendental reflection pertains to the *nexus* of matter and form of one's doctrine.

For the present undertaking, Critical doctrine of method seems to demand that attention be paid to what William James called the penultimate problem in understanding mind-brain, viz. its *statement* problem:

To state [the mind-brain problem] in elementary form one must reduce it to its lowest terms and know which mental fact and which cerebral fact are, so to speak, in immediate juxtaposition. We must find the minimal mental fact whose being reposes directly on a brain-fact; and we must similarly find the minimal brain-event which will have a mental counterpart at all. [James, *Principles of Psychology*, vol. I, pg. 177]

While this seems likely to be true within some context of understanding, the problem with James' statement is the obscurity of such ideas as "minimum mental fact," "brain-fact," and "brain-event." He is at some level talking about object-to-object comparison. Critical canon, though, cautions that as we begin to consider James' problem we understand such comparisons as the joint actions of the *Verstandes-Actus* of *Comparison* and *Reflexion* in sensibility (Wells, 2009). Otherwise we will immediately run into the problem of naming a *standard* of comparison, without which no objectively valid comparison is possible at all. *Comparison* is the synthesis of a mathematical equivalence relation and *Reflexion* is the synthesis of a mathematical congruence relation. Both types of relations are required if one is to judge that "the being of a mental fact reposes directly on a brain-fact" regardless of whatever these two things might be.

Mental physics says as much as this when we speak of the overlap of principal quantities in mathematics with observables in physical Nature. We cannot say that some fact of physical Nature (facet A) and some fact of mathematical Nature (facet B) correspond to form a theoretical context at all, as illustrated in figure 1, without an objectively valid real comparison.

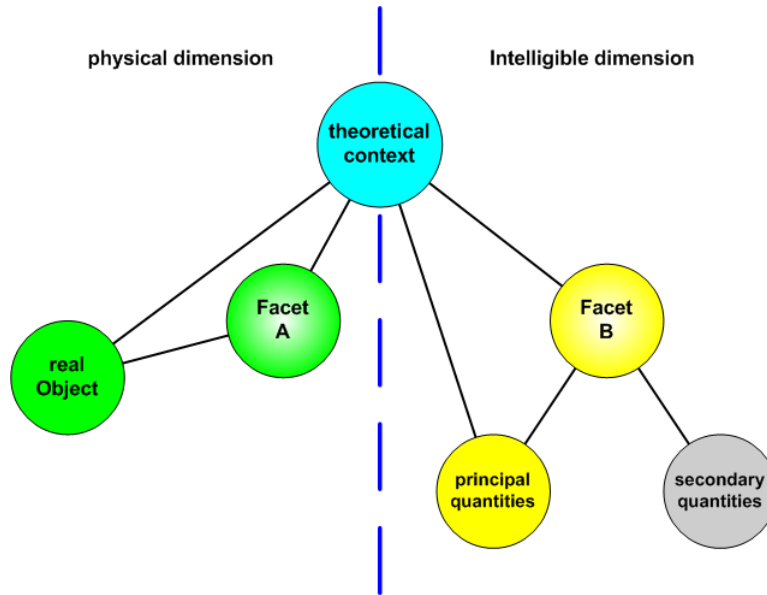


Figure 1: The unification of physical Nature and mathematical explanation by theoretical context.

The immediately foremost issue with the problem of comparison is that the *Verstandes-Actus* of the synthesis in sensibility are entirely non-objective, producing both affective perceptions and intuitions. All the outcomes of the *Verstandes-Actus* are *judged* by the process of reflective judgment, which obeys as its fundamental acroam the *subjective* acroam of formal expedience. The lesson Critical epistemology holds for science is this: *all scientific doctrines are uncertain at some degree of holding-to-be-true vs. holding-to-be-false*. The set-membership methodology of a Critical science explicitly recognizes this *psychological* Nature of science. *Discipline* in using the *Verstandes-Actus* of abstraction, which must be carefully taught and developed through educational experience, requires we develop an applied canon by which maxims of judgmentation (in the manifold of rules) are developed for regulating the use of reasoning and judgmentation in the construction of theoretical contexts.

The historical evidence of the real *Existenz* of this issue in science is revealed by the classical controversies scientists have found themselves engaged with in regard to various methodologies for establishing or refuting scientific theories. Examples of this include the falsificationists' doctrine, the justificationists' doctrine, and the doctrine of probabilism. Even set membership doctrine faces this issue because at some point in its methodology it must *assign* some level of distinguishability and this assigned level has the formal consequence of fixing the cardinality of the set of mathematical solutions that share the common property of being indistinguishable in relationship to all currently known empirical data and all *a priori* knowledge of the Nature of the system¹ to the study of which the methodology is being applied. At present, set membership theory possesses no disciplinary or canonical rules for objectively fixing this level of distinguishability, nor do its practitioners agree on a common conventional method for doing so.

II. Set Membership Theory and the Appraisal Problem

Receptivity presents the human being with no pre-fixed knowledge of any "order in Nature."

¹ In this particular instance, *a priori* knowledge is not transcendental 'know-how' knowledge but, rather, is the body of belief-concepts underlying the *mathematical* representation of the object of study. Belief-concepts are formulated under the transcendental law of the Axioms of Intuition and serve as axioms in the process of thinking. Belief is *unquestioned* holding-to-be-true in judgmentation (Wells, 2009).

That we understand Nature systematically, i.e. that we *think* order-in-Nature, is foundationally due to the process of teleological reflective judgment, which is tasked, among other functions, with organizing the representation of order-in-Nature (Wells, 2006). When scientific knowledge becomes sufficiently erudite for scientists to recognize the so-called "verificationist vs. falsificationist" problem and to seek standards for normative conventions in judging scientific theories, they then are confronted by the mental physics of the process of pure practical Reason – the Nature of which is *impatient*. Reason knows no objects and feels no feelings. Its sole concern is practical equilibration in the Organized Being, and through ratio-expression it seeks the most direct route to achieving this. This is what Kant meant when he wrote of the propensity (*Hang*) of the process of speculative Reason to produce transcendent ideas beyond the horizon of possible human experience (Kant, *Critique of Pure Reason*).

Thus, we are faced from the outset by a requirement to deal with an even more fundamental problem than James' penultimate problem. One philosopher of mathematics who recognized this issue was Imre Lakatos. Although there are many flaws in Lakatos' system of metaphysics², he was able to offer in outline form a cogent insight with regard to this issue. Programs of research methodology, he wrote, must recognize that the problem of verification vs. falsification of scientific doctrines is a *historical* problem – by which I mean a problem to be resolved by a *social-natural* science of history³. He wrote,

I am going to propose a new theory of how to appraise such methodologies of science . . . I shall show that methodologies may be criticized without any direct reference to any epistemological (or even logical) theory, and without using directly any logico-epistemological criticism. The basic idea of this criticism is that *all methodologies function as historio-graphical (or meta-historical) theories (or research programs) and can be criticized by criticizing the rational historical reconstructions to which they lead.* [Lakatos, *Methodology*, pg. 122]

I comment without polemic that "without direct reference" does not mean "without reference" to epistemology⁴, that criticism helps provide discipline but establishes no canon, and that Lakatos does not adequately survey the state of history as a science. He was no positivist, but he was something of a "neo-Eclectic empiricist" in his metaphysics of mathematics.

What I do wish to emphasize is that Lakatos touches upon an important point by bringing history into the context of scientific theorizing. As noted earlier, a history-of-pure-Reason is regarded by Kant as an important (albeit unfinished) part of a transcendental doctrine of method. It serves the Modality function in scientific judgment, and such a function is always a judgment *of* a judgment, not of the *object* of that judgment. The Modality function adds nothing to our knowledge of the object as object but does provide the connection between the scientist and that-which-he-is-thinking-about. Kant called this the *metaphysical nexus* in judgment⁵.

The transcendental requirement for this component of scientific methodology is made clear by the earlier observation in regard to the subjectivity of the *Verstandes-Actus* within the synthesis in sensibility. It is no mere coincidence that the circumstance with which we have to deal lies right

² It is not unfair to Lakatos – or, at least, not entirely unfair – to characterize his metaphysics as neo-Eclecticism. One finds in his works very close similarities with the philosophical system Cicero claimed, in *Tusculan Disputations*, was the one to which he personally subscribed.

³ I will at once point out that this, too, faces a problem in its practical implementation. History is not at present a social-natural science but, instead, as a science is currently a mere historical doctrine of Nature under Kant's taxonomy of sciences. There is a great deal of foundational work yet to be done to turn history into a proper social-natural science.

⁴ No proper natural science can be built *without* a foundation in Critical epistemology.

⁵ *Kritik der reinen Vernunft*, B: 201-202 fn. (3: 148-149 fn.).

at the junction of facets A and B in Slepian dimensioning (figure 1) because the mathematical ideas of congruence relations and equivalence relations are intimately linked to the synthesis of intuitions. (Intuitions are axioms produced in the process of thinking by the free play of imagination and understanding). Just as intuitions (and the initial belief-concepts their transformation in the synthesis of re-cognition produces) are subject to later questioning, and a consequential re-structuring of the manifold of concepts, so also is set membership-based appraisal of scientific models (theories) subject to this same questioning-of-belief that arises from the process of aesthetical reflective judgment. The core of Lakatos' idea *is sanctioned by mental physics as an epistemological necessitation*.

Kant noted,

All Knowledge⁶ is either empirical, i.e., derived from experience, or rational: arising from reason, hence possible *a priori* and self-supporting. Among the former will have been counted experience proper and history (i.e., reliable reports, hence Knowledge from the experience of others). The second kind of certitude is independent from all experience.

All empirical certitude is combined with consciousness of the contingency of the truth; for experience teaches well that something is constituted in one way or another or that something has happened, but never teaches that it could not have been constituted or happened otherwise. [Kant (c. 1783-84), *Reflexionen zur Metaphysik*, 18: 290]

He further remarked that

The utility-aim of philosophical history subsists in the preparation of good models, and the presentation of instructive mistreatments likewise, in the knowledge of the natural progress of reason from ignorance (not crude error) to knowledge. [*ibid.* (c. 1776-78), 18: 12]

Kant's examples of "instructive mistreatments" (*lehrreicher Vergehungen*) were recapitulations of philosophy or science (natural philosophy) doctrines that were either failed or unsound systems. He would then append to them a Critical analysis of where and how these doctrines had erred and prejudicially "mistreated" metaphysics or science. By "philosophical history" we understand him to mean the employment of history in metaphysics as part of a doctrine of method (rather than as either history-of-philosophy or philosophy-of-history). Unfortunately, Kant never did complete the "filling in" of the history component of the Critical doctrine of method. The one essay he did produce on history was a minor work⁷ that must be called a mere romantic speculation with pre-Hegelian overtones that does more to expose some of Kant's personal prejudices than to make any contribution to history as a social-natural science.

Now, any mathematical theory of an object in Nature is, by virtue of mathematical concepts being the representations of *noumena*, a theoretical *model* of that object and produces one or more principal quantities associated with sensible phenomena. Critical epistemology, however, teaches us that valid application of association between a mathematical principal quantity and concepts of a sensible phenomenon *must* result in a determinant judgment of understanding that is bound by a very specific *momentum* of Quality. Specifically, it is one in which the judgment of relationship between principal quantity and phenomenal concept has the *category of limitation* as its

⁶ *Wissen*, i.e., systematic and unalterable assertion of truth with consciousness that holding-to-be-true is grounded in judgments that have apodictic Modality with both objectively and subjectively sufficient grounds of understanding. Capitalization (Knowledge) is used to distinguish the translation of *Wissen* from *Erkenntnis* (knowledge, i.e., any conscious representation or capacity for making such a representation by or through which meanings are determined; used in the narrow sense, knowledge is a cognition held-to-be an alterable assertion of truth).

⁷ Kant (1784), *Idee zu einer allgemeinen Geschichte in weltbürgerlicher Absicht*, in *Kant's gesammelte Schriften, Band VIII*, Berlin: Walter de Gruyter & Co., 1923.

momentum of Quality in the judgment. This has a direct bearing on the proper context in which set membership methodology is to be interpreted.

We must regard the set membership methodology as a method that uses its solution set as a statement of *what is not-rejected*. Set membership's "set of consistent solutions" are not to be seen as being consistent because they are not rejected but, rather, they are not rejected because they have not been shown to be inconsistent with all current empirical data and all *a priori* knowledge of the system being modeled. Consistent solutions are the not-rejected solutions. The members of the solution set are in some sense like Lakatos' idea of a "series of theories" (*Methodology*, pg. 34) except we have here members of a disjunction rather than a series in Relation. Lakatos' idea of an "appraisal of a series of theories" is in many ways similar to set membership's determination of the possible-solutions set.

Lakatos proposed these ideas from critiquing the historical behavior of scientists on the issue of when they will reject an established theory that appears to have been contradicted by some new fact *vs.* when they will continue to accept and continue to use an established theory even when faced with some apparently contradictory finding or findings. He wrote,

Let us say that . . . a series of theories is *theoretically progressive* (or '*constitutes a theoretically progressive problem shift*') if each new theory has some excess empirical content over its predecessor, that is, if it predicts some novel, hitherto unexpected fact. Let us say that a theoretically progressive series of theories is also *empirically progressive* (or '*constitutes an empirically progressive problem shift*') if some of this excess empirical content is also corroborated, that is, if each new theory leads us to an actual discovery of some *new fact*. Finally, let us call a problem shift *progressive* if it is both theoretically and empirically progressive, and *degenerating* if it is not. We '*accept*' problem shifts as '*scientific*' only if they are at least theoretically progressive; if they are not, we '*reject*' them as '*pseudo-scientific*'. Progress is measured by the degree to which a problem shift is progressive, by the degree to which the series of theories leads us to the discovery of novel facts. We regard a theory as '*falsified*' when it is superseded by a theory with higher corroborated content.

This demarcation between progressive and degenerating problem shifts sheds new light on the appraisal of *scientific – or, rather, progressive – explanations*. If we put forward a theory to resolve a contradiction between a previous theory and a counterexample in such a way that the new theory, instead of offering a content-increasing (scientific) *explanation*, only offers a content-decreasing (linguistic) *reinterpretation*, the contradiction is resolved in a merely semantical, unscientific way. *A given fact is explained scientifically only if a new fact is also explained with it.*

Sophisticated falsification thus shifts the problem of how to appraise *theories* to the problem of how to appraise *series of theories*. Not an isolated *theory*, but only a series of theories can be said to be scientific or unscientific; to apply the term '*scientific*' to one *single* theory is a category mistake.

The time-honored empirical criterion for a satisfactory theory was agreement with the observed facts. Our empirical criterion for a series of theories is that it should produce new facts. *The idea of growth and the concept of empirical character are soldered into one.*

This revised form of methodological falsification has many new features. First, it denies that "in the case of a scientific theory, our decision depends upon the results of experiments. If these confirm the theory, we may accept it until we find a better one. If they contradict the theory, we reject it." It denies that "what ultimately decides the fate of a theory is the result of a test, *i.e.* an agreement about basic statements." Contrary to naive falsificationism, *no experiment, experimental report, observation statement, or well-corroborated low-level falsifying hypothesis alone can lead to falsification. There is no falsification before the emergence of a better theory.* . . . Falsification can thus be said to have a '*historical character*'. [Lakatos, *Methodology*, pp. 33-35]

What, though, does Lakatos mean by "series of theories"? His treatise is obscure on this point. We cannot, for example, take this to mean phlogiston chemistry vs. modern chemistry or caloric thermodynamics vs. modern thermodynamics vs. statistical mechanics. Too many present day scientists use the word "theory" too loosely. It is important to understand that a theory is not a fact; theories are proposed explanations of facts and are properly judged in terms of how well, or not, the explanation sets the facts to be explained in context with other facts in the overall structure of Nature. In the youth of modern science, what is usually called a "theory" today was called an hypothesis, a terminology that recognized the open-ended nature of knowledge discovery. Newton set down as a "rule of reasoning in [natural] philosophy" that

*In experimental philosophy we are to look upon propositions inferred by general induction from phenomena as accurately or very nearly true, notwithstanding any contrary hypothesis that may be imagined, till such time as other phenomena occur, by which they may either be made more accurate or liable to exceptions. This rule we must follow, that the argument of induction may not be evaded by hypothesis. [Newton, *Mathematical Principles of Natural Philosophy*, Bk. III]*

If we are to speak of a "series of theories," this manner of speaking can only refer to some collection of contending hypotheses that have been put forward over time seeking to explain the same body of facts. Newton, following Francis Bacon's earlier prescription in *Novum Organum*, held that induction was the only reliable method for developing scientific explanations. Both men, however, failed to understand that inferences of induction are the product of teleological reflective judgment and are *subjective* rather than objective judgments. Modern mathematicians, too, are usually hostile to any suggestion that mathematical induction is subjective rather than objective; this has been the prevailing attitude throughout much of the twentieth century to the present day. One critic of this dogma was the renowned mathematician Henri Poincaré:

We cannot therefore escape the conclusion that the rule of reasoning by recurrence is irreducible to the principle of contradiction. . . This rule, inaccessible to analytical proof and to experiment, is the exact type of the *a priori* synthetic intuition. . .

Why, then, is this view imposed upon us with such an irresistible weight of evidence? It is because it is only the affirmation of the power of the mind which knows it can conceive of the indefinite repetition of the same act when the act is once possible. The mind has a direct intuition of this power, and experiment can only be for it an opportunity of using it, and thereby becoming conscious of it. . . Induction applied to the physical sciences is always uncertain, because it is based on the belief in a general order of the universe, an order which is external to us. Mathematical induction – i.e., proof by recurrence – is, on the contrary, necessarily imposed on us, because it is only the affirmation of a property of the mind itself. [Poincaré, *Science and Hypothesis*, chap. 1]

Mental physics tells us Poincaré is correct about this except for a few minor issues of semantics.

Appraisal of scientific theories obviously speaks to the metaphysical *nexus* of Modality in scientific methodology (i.e.: it *might-be* true/false; it *is* true/false; it *must-be* true/false). Thus Lakatos appears to stand in agreement with Kant on this point that history is a required part of the overall doctrine of method and its role is that upon which are based the judgment of hypotheses. As hypotheses ("theories") are themselves representations of understanding, the role of history in doctrine of method is properly a role in the Modality of method.⁸

However, this methodological picture is not a precise description of what is currently practiced in set membership theory. In set membership theory, a model does not make a single *point* "prediction" (model result) but, instead, produces a *set* of "consistent" results. This is

⁸ Lakatos further elaborated his idea in his *Mathematics, Science, and Epistemology*.

termed the *solution set* of the model. This terminology means that a measurement or observation that falls anywhere within the range of this set of results is *regarded* as consistent with the mathematical model's principal quantity. The theory is said to be "consistent with the facts." *Every mathematical model that produces the same solution set is likewise said to be "consistent with the facts"* and so it follows from Slepian's principle that methodology acknowledges as actual a hypothetical ensemble of "equally consistent" *possible* theories and regards all members of this ensemble as *the un-rejected set of mathematical theories in facet B*. What we have could not be called a *series*, in the context of a temporal sequence of scientific theories, but, rather, a set of *simultaneous* possible theories arranged together under a single theory-Object⁹.

The concept structure involved here is a Critical disjunctive proposition (Wells, 2009: chapter 6, §3.4). The members of the disjunctive structure are co-determining inasmuch as any determination made on one member is at the same time a reciprocal determination of all the other members. However, Critical epistemology tells us that the logical nature of Critical disjunction is not as simple as the logical disjunction ('OR') presented in either classical or symbolic logic. Determinant judgments of disjunction do not formally operate on single concepts but, rather, on entire spheres of concepts. Furthermore, the temporal sequence in which these concept structures are formed affects subsequent judgments, a dynamical factor that is altogether left out of both classical and mathematical logic (Wells, *ibid.*). When, then, we move to consider Critical set membership mathematics and how new experiences alter set membership solution sets, we must do so from a basis in Critical epistemology rather than as set membership formalism is presently set up. Here the doctrine of history in Critical methodology will have to come into sharper focus because this Modal element speaks to the proper treatment of what is conventionally called the "error bound" in set membership formalism. The "historicism" of Lakatos' idea is one proposal for how to properly handle the Modal judgment (which, in effect, is the role approximated by the error bound parameter in a set membership model).

III. Slepian's Principle

Set membership theory is the name given to a family of related mathematical methodologies. Its importance in Critical methodology arises from the linkage it provides between the intelligible world of mathematics and the phenomenal world of physical Nature. The fundamental principle of this linkage is called Slepian's principle (Wells, 2009).

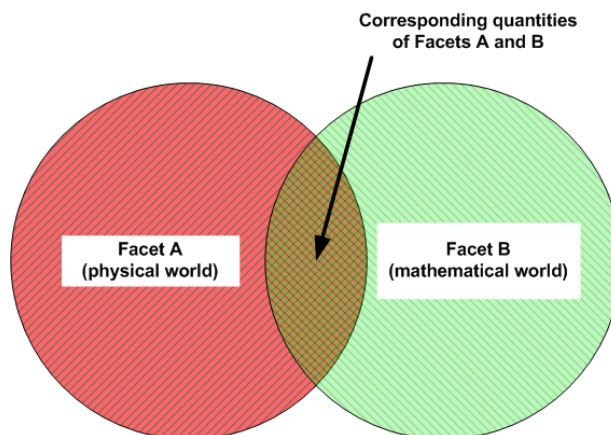


Figure 2: Slepian's model of physico-mathematical correspondence.

⁹ I stress the word "possible" here because not every theory that *could* belong to this ensemble has necessarily already been posited. New theories (models) can be added and old theories removed from the ensemble as experience brings forth more phenomena the theory-ensemble is required to explain.

Slepian's principle, first enounced by information theorist David Slepian in his Shannon Lecture at the 1975 Annual Meeting of the Information Theory Society, is the Critical solution to a centuries old issue: How is it possible for mathematics (which is so clearly the product of human intellect) to make true and objectively valid propositions concerning physical phenomena (which, we presume, do not have their origins from human intellect)? Slepian's solution is a *canon of methodology* in applying mathematical reasoning to the understanding of Nature. We have here to deal, Slepian said, with two "worlds" – that of sensible physical Nature (facet A) and that of supersensible (*noumenal*) mathematical Nature (facet B) – and we must recognize that there are aspects in each facet that have no correspondence at all with any object in the other facet. On the other hand, there are objects in each facet that can be placed in a *practical* relationship with one another, and that this pairing of (facet A Object, facet B Object) constitutes a practical overlap or "intersect" *jointly definable* for these two "worlds." Figure 2 illustrates Slepian's "two worlds" model of physico-mathematical correspondence.

Slepian's solution is a canon because it does not deal with objects but rather with joint definition of the *combination* of Objects, one placed in facet A, the other in facet B. His principle is a principle for epistemological requirements that must be imposed upon all mathematical reasoning insofar as that reasoning can be objectively valid. Objects of mathematics are without exception *noumena* and as such can never be *immediately* presented in possible experience in the *context* of physical Nature. Their *real* context is intelligible rather than physical. But, as Slepian showed, it is possible through methodology to place some mathematical Objects in a relationship with physical Objects in such a way that what the relationship states is neither less nor more than what is accessible in sensuous experience. Mathematical Objects that can be so placed are called the *principal quantities* of mathematics. Mathematical Objects for which such a placement is not possible are called the *secondary quantities* of mathematics. The latter stand outside the overlap depicted in figure 2 and are properly regarded as, to use a metaphorical phrase, "orthogonal to the plane of physical Nature."

Principal quantities, by contrast, stand at the very edge of the horizon of possible experience where they are *united*, by *theoretical* context, with ideas of *noumena* that represent the highest level of *objectively valid* physical understanding. In Critical epistemology, the *Dasein* of such an object is held-to-be-necessary for the possibility of experience as human beings come to have experience. But although we have knowledge of the object's *Dasein* in facet A, we can have utterly no objectively valid knowledge of the object's *Existenz* in facet A. Principal quantities are ideas of the object's *Existenz* in *facet B*. Such a combination cannot be justified with objective validity in any ontology-centered system of metaphysics, but can be and is justified by epistemology-centered Critical metaphysics. This follows from the epistemological understanding that ***all objects are real in some contexts, unreal in other contexts, and non-real in yet other contexts.*** The ghost of Hamlet's father is real in the context of the play *Hamlet*, unreal in the context of being a thing that actually haunts anyone in Denmark, and is non-real (has no context at all) in economic theory or Boolean algebra. An object is real when one has a concept of the object connected by determinant judgments with other concepts, that give it context and meaning, and that has in this context some connection to at least one actual sensuous experience.

To understand *how* it is real is to understand its *Existenz*, and for objectively-valid *noumena* this understanding is by means of secondary quantities. These entities of pure mathematics serve a practical purpose we would not falsely describe by saying that they bring continuity to the "surface" of the horizon of possible experience. Figure 3 illustrates what I mean by this. To employ a metaphor, they provide the "surface tension" that "holds physical Nature together" in human understanding. To further perfect this understanding is, of course, the task of science proper. Slepian's principle grounds a canon of physico-mathematical reasoning by which science can accomplish its task with objective validity.

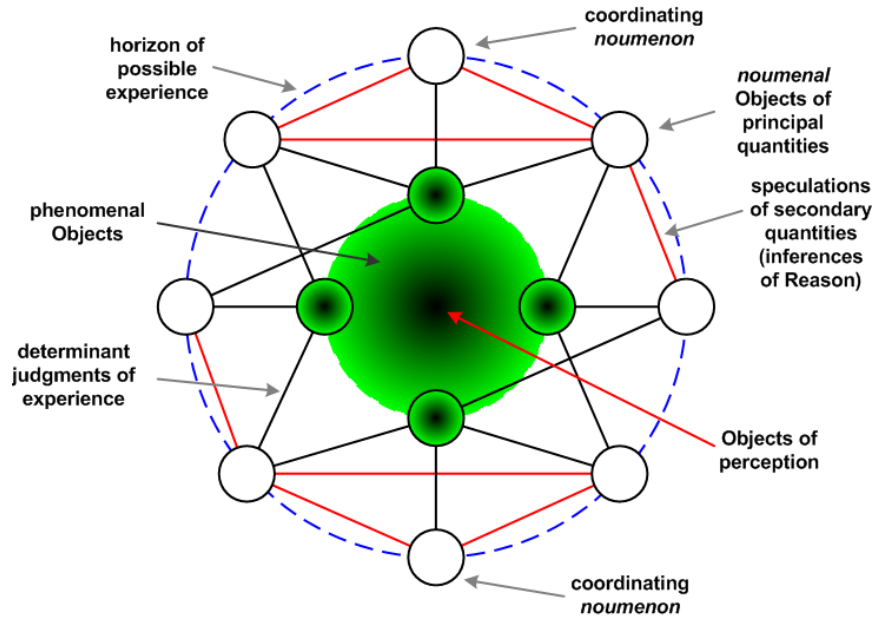


Figure 3: Illustration of the structure of Critical scientific ontology.

It is accurate to say that Slepian introduced and illustrated his principle by applying it to one specific problem of long-standing interest in the science of system theory (called the bandwidth paradox). It is also accurate to say that he sketched out his canon but his Shannon Lecture did not go so far as to provide a doctrine for it. Nonetheless, its significance was well enough appreciated at the time that the unusual step was taken of having his lecture published verbatim in *The Proceedings of the IEEE*¹⁰ the next year (1976). Taken no further than where Slepian left it in 1976, its statement would have had little broad utility and been little more than a call for further action. As it happened, though, there was already at that time a nascent doctrine being slowly put together that, upon examination, is architectonic and founds a discipline for applying Slepian's principle to all topics of science. That doctrine is set membership theory.

IV. Slepian's Principle and Set Membership Theory

Any physical theory is a model of Nature and attempts to relate some set of input factors corresponding to objects of facet A to some set of output factors corresponding to objects of facet A. Let us call a theory *T* a *quantitative* theory if its set of output factors $[x_1, x_2, \dots, x_n]$ can all be unambiguously assigned specific quantities once all its input factors $[w_1, w_2, \dots, w_j]$ have also been similarly determined. The specific quantities so assigned are determinations of the magnitudes of the factors, and in a mathematical theory these quantities are usually called numbers¹¹. We will call two theories, T_1 and T_2 , *comparable* if: (1) each contains within its output factors some *common subset* of results $\mathbf{X} = [x_1, x_2, \dots, x_N]$ corresponding to the same objects in

¹⁰ *The Proceedings* is the most prestigious technical journal published by the Institute of Electrical and Electronic Engineers. Most papers appearing in it are by invitation of the Institute.

¹¹ The idea of "number" in mathematics is very general and leads to classifications of different types of numbers, e.g., natural numbers, integers, real numbers, complex numbers, etc. All that is necessary for our discussion here is that, whatever measures of magnitude are used for determining each specific factor in a theory, they are such that any two numbers that may be used to determine the same factor can have some measure of the difference between them specified by some mathematical metric function when each number is regarded as a point in a metric space. By "metric function," "metric space" and "point" I mean the standard mathematics definitions provided in Nelson's *Dictionary of Mathematics*.

facet A; (2) both theories can have their sets of input factors completely determined for each trial of the theory; (3) the various sets of numbers used to determine factors in the theory¹² can be placed in some mathematical order relation such that the concepts "less than" and "greater than" can be applied to their relative numerical values; and (4) all results \mathbf{X} can also be placed in such a mathematical order relation. A *trial* of a theory is the set $\mathbf{X} = [x_1, x_2, \dots, x_N]$ of common comparable output factors the theory produces for a particular determination of its input factors. We will call \mathbf{X} the *solution* of the theory for a particular trial. Any specifically determined \mathbf{X} , i.e. an \mathbf{X} for which numbers have been determined for all its factors, is called the *quantity* of the result. Because all the factors in an \mathbf{X} are to correspond to objects in facet A, the mathematical object represented by an \mathbf{X} is a *principal quantity* of the theory. Because of condition (4) above, the set of possible quantities \mathbf{X} constitutes a mathematical N-dimensional metric space (sometimes called a "hyperspace" when $N > 3$). If we denote the metric function measuring the difference between two quantities \mathbf{X}_1 and \mathbf{X}_2 by $\rho(\mathbf{X}_1, \mathbf{X}_2) \geq 0$, we will call the specific outcome of applying $\rho(\mathbf{X}_1, \mathbf{X}_2)$ to these quantities the *distinction* between \mathbf{X}_1 and \mathbf{X}_2 .

Now, principal quantities are said to be empirically determined because in order to make a determination of \mathbf{X} the input factors $[w_1, w_2, \dots, w_j]$ must be measured (by observation and often by means of measuring instruments) so that numbers can be assigned to each factor. However, all such measurements are never more than appearances of the facet-A objects to which the factors are to correspond – and, hence, quantities $[w_1, w_2, \dots, w_j]$ are likewise principal quantities. Slepian pointed out that all such empirical determinations can be quantified only to some finite level of accuracy and precision and, therefore, some degree of uncertainty is always inherent in any determination of a quantity \mathbf{W} or \mathbf{X} . Mathematically, this means there is always some number $\varepsilon > 0$ below which a distinction $\rho(\mathbf{X}_1, \mathbf{X}_2)$ is no longer an object of any *actual* experience. For any empirical $\rho(\mathbf{X}_1, \mathbf{X}_2) < \varepsilon$, the distinction as an object passes beyond the horizon of possible experience, is no longer part of *empirical* Nature, and the distinction becomes a *secondary* quantity of mathematical facet B. \mathbf{X}_1 and \mathbf{X}_2 are then said to be *rationally* distinct but *empirically* indistinct. We may call ε the *empirical uncertainty* of the theory.

Such uncertainty also attends the ordering of solutions \mathbf{X}_1 and \mathbf{X}_2 produced by different theories T_1 and T_2 . Slepian calls two theories such that $\rho(\mathbf{X}_1, \mathbf{X}_2) < \varepsilon$ *indistinguishable at level ε* by the particular metric function $\rho(\mathbf{X}_1, \mathbf{X}_2)$. Here we would call ε *natural empirical uncertainty* because it will be determined by the actual and practical capacity to determine $[w_1, w_2, \dots, w_j]$. **Slepian's principle** states *if the members of a set of theories all produce principal quantities that are indistinguishable at level ε according to some criterion of distinguishability then these theories are empirically equivalent at level ε* . The set of all such empirically equivalent solutions is called the *solution set*. Figure 4 illustrates this idea for $N = 2$ output factors.

Slepian correctly pointed out that an observer or an experimenter possesses no information by which he can make any objectively valid proposition that would distinguish one member of the solution set from any other member. Any *paradoxes* ascribed to one theory but not encountered in another will be found to originate from one or more invalid propositions the theorist might have proposed for the theory. In Kantian terminology such a proposition must be called a proposition posited of a *Ding an sich selbst*, i.e., a *noumenon* beyond the horizon of possible experience. Such a noumenon is called a *thing-as-we-cannot-know-it* and the proposition itself is *formally undecidable*, i.e., can be called neither true nor false.

¹² It is not necessary that the numbers used to measure factors in the theory all be drawn from the same number set for each factor in the theory. All that we must require is that all numbers used as measures of the *same* factor are drawn from the same number set. Thus, within a factor set we might, e.g., have x_1 be drawn from the set of integers, x_2 from the set of complex numbers, x_3 from the set of binary digits, etc.

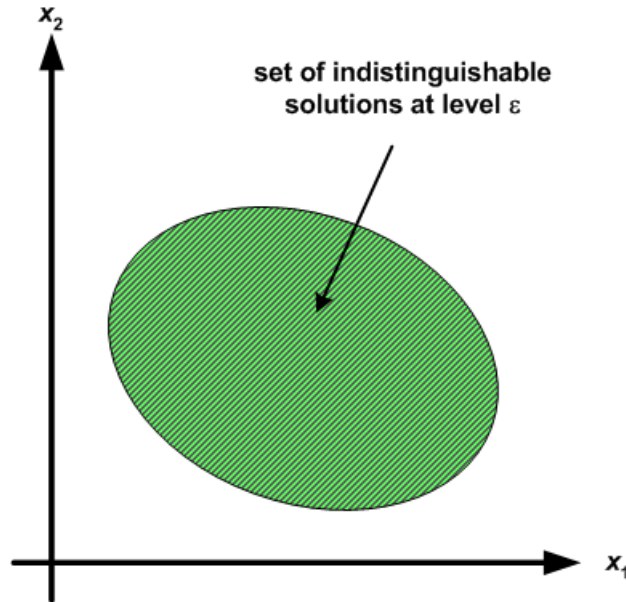


Figure 4: Illustration of a set of empirically equivalent solutions according to Slepian's principle

Slepian qualified his principle by saying that the metric function $\rho(\mathbf{X}_1, \mathbf{X}_2)$ is to be defined according to some "criterion of distinguishability." The one he employed in his Shannon Lecture belongs to a class of metric functions that system theorists usually call an *energy function*. Here the term "energy" is an idea of an Object of facet B under which one finds a great diversity of specific "types" or "kinds" of "energies." Physics and chemistry, in practice, make extensive use of this idea but here it is important to keep in mind that "energy" is a *noumenon* and is not itself the object of any possible immediate sensuous experience. Nobel laureate Richard Feynman explained the practical meaning physicists give this term in the following way:

There is a fact, or if you wish, a *law*, governing all natural phenomena that are known to date. There is no known exception to this law – it is exact so far as we know. The law is called the *conservation of energy*. It states that there is a certain quantity, which we call energy, that does not change in the manifold of changes the universe undergoes. That is a most abstract idea, because it is a mathematical principle; it says there is a numerical quantity which does not change when something happens. It is not a description of a mechanism, or anything concrete; it is just a strange fact that we can calculate some number and when we finish watching nature go through her tricks and calculate the number again, it is the same. (Something like the bishop on a red square [in chess], and after a number of moves – details unknown – it is still on some red square. It is a law of this nature). . .

First, when we are calculating the energy, sometimes some of it leaves the system and goes away, or sometimes some comes in. In order to verify the conservation of energy, we must be careful we have not put any in or taken any out. Second, the energy has a large number of *different forms*, and there is a formula for each one. These are: gravitational energy, kinetic energy, heat energy, elastic energy, electrical energy, chemical energy, radiant energy, nuclear energy, mass energy. . .

It is important to realize that in physics today, we have no knowledge of what energy is. We do not have a picture that energy comes in little blobs of a definite amount. It is not that way. However, there are formulas for calculating some numerical quantity, and when we add it all together it gives "28" – always the same number. It is an abstract thing in that it does not tell us the mechanism or the reasons for the various formulas. [Feynman, *The Feynman Lectures on Physics*, vol. I, chap. 4]

Different criteria will generally result in different definitions of metric functions and this raises an important subjective factor, namely that of choosing the $\rho(\mathbf{X}_1, \mathbf{X}_2)$ function, that must be taken into account in the development of a canon in a doctrine of method. This is the comparison issue, raised in sec. I, merely transplanted to the level of conventional practices a science adopts. Slepian stated an important criterion for the establishment of criteria and their applications: *The secondary quantities of a theory, being unobservable in facet A, must have their linkage to principal quantities (in facet B) be such that the determinations of a principal quantity \mathbf{X} are insensitive to changes in the numerical determinations of secondary quantities.* He remarked,

One can, of course, consider and study any model one chooses to. It is my contention, however, that a necessary and important condition for a model to be *useful* in science is that the *principal quantities of the model be insensitive to small changes in the secondary quantities.* Most of us would treat with great suspicion a model that predicts stable flight for an airplane if some parameter is irrational but predicts disaster if that parameter is a nearby rational number. Few of us would board a plane designed from such a model. [Slepian, "On bandwidth."]

Examples of Slepian's principle can be found elsewhere in science and mathematics and these examples are attention-deserving because they arose independently of Slepian's work. One of these is found in the practice of renormalization in physics' theory of quantum electrodynamics. Here it is found that the determinations of such principal quantities as charge and mass can be computed without significant change in determination from a very wide range of possible secondary quantity determinations. A second example is provided by Robinson's theory of non-standard analysis in mathematics, where his technique can be interpreted in terms of Slepian's principal and secondary quantities. In non-standard analysis, these are defined in terms of what are called "the standard universe" and "the non-standard universe" with corresponding elements called "reals" and "pseudo-reals." Non-standard analysis makes the mathematical treatment of infinitesimals formally precise and resolves the old controversy between Berkeley and Newton.

Slepian's principle is applicable to single trials of a theory and to single trial comparisons of different mathematical theories. However, the principle by itself does not address the larger issue of falsification in Lakatos' sense, which involves multiple trials. For that we must turn to set membership theory proper.

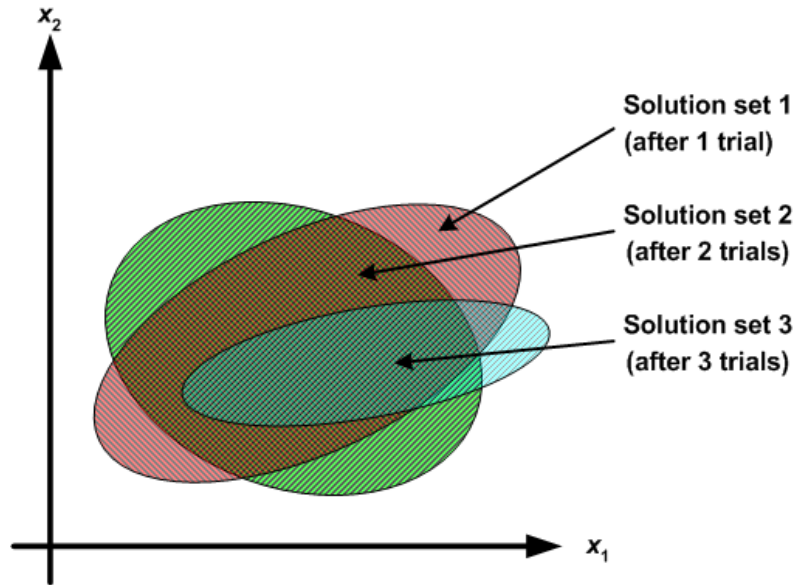


Figure 5: Illustration of successive determinations of the solution set over multiple trials.

Let models M_1, M_2, \dots, M_M (or theories T_1, T_2, \dots, T_M) produce results X_1, X_2, \dots, X_M , respectively, for the same input set W_1 . Furthermore, let these results be indistinguishable at some level ε . The set union of these results then constitutes a solution set $\Xi_1 = \{ X_1, X_2, \dots, X_M \}$ such as illustrated by figure 4 or by the red-shaded solution set in figure 5. Now let a new input set W_2 be applied to these same models (or theories), resulting in another solution set Ξ_2 . For purposes of illustration, suppose this is represented by the green-shaded solution set in figure 5. Note that if each specific model (or theory) produces only one result X_j for each input W then Ξ_2 permits inclusion of some models (or theories) that were not included in Ξ_1 , contains some of the same models (or theories) as contained in Ξ_1 , and excludes some of the models (or theories) that were contained in Ξ_1 . The set of models contained in the solution sets for both trials, $\Xi^{(2)} = \Xi_1 \cap \Xi_2$, is the set that is consistent with both trials and thus constitutes an *historical record* of trial results up to this point. This is illustrated in figure 5 by the shaded area where the two solution sets overlap. $\Xi^{(2)}$ therefore represents the *empirically consistent solution set* up through the second trial because it contains all those models (or theories) that have produced results consistent with all *a priori* knowledge of the system plus all currently known empirical data.

Now let there be a third trial, W_3 , that produces a solution set Ξ_3 (illustrated by the light blue set in figure 5). The empirically consistent solution set after the third trial is $\Xi^{(3)} = \Xi^{(2)} \cap \Xi_3$. Figure 5 illustrates this set intersect as well. This procedure can be continued in an unlimited series of trials with the outcome that after the n th trial the empirically consistent solution set is the set defined by $\Xi^{(n)} = \Xi^{(n-1)} \cap \Xi_n$. This is the operational definition of the set membership method.

In a succession of trials there are two logical limiting cases in regard to solution set $\Xi^{(n)}$. The first is that the solution set may be unchanged, i.e., $\Xi^{(n)} = \Xi^{(n-1)}$. This occurs if every model (or theory) already subsumed under $\Xi^{(n-1)}$ produces a result X already contained in $\Xi^{(n-1)}$. This is possible because of Slepian's empirical uncertainty factor $\varepsilon > 0$. System theorists typically call ε the *error bound* of the system and it is used to establish what is usually termed the *cut-sets* of the system¹³. This outcome means that trial W_n provided no new information about the Nature of the system, as discussed in Fogel and Huang (1982). This case is the limiting case encountered in practice if: (1) no instrumental or other improvements are developed that permit a reduction in the empirical uncertainty, ε ; and (2) the *a priori* knowledge thought to be true of the natural system contains no errors leading to paradox, paralogism, or antinomy¹⁴ in the model or theory.

The second limiting case is that in which eventually some trial W_n produces no result contained within any subset of $\Xi^{(n-1)}$, i.e., $\Xi^{(n-1)} \cap \Xi_n = \emptyset$, the empty set. This case occurs if there is some fundamental experimental or observational error (W_n is incorrectly determined during some one or more trials) or if the accepted basis B in *a priori* knowledge of the system omits something essential to one's understanding of the system or introduces something false into its Object. It is possible for B to contain inessential knowledge (whether true or false) without this outcome resulting, but the null result will occur if B contains essentially false knowledge. This has been demonstrated by McCarthy and Wells (1997).

There is also another way for the second limiting case to occur. This is for the system itself to alter or change in some way during the series of trials. Such a system can generally be called a

¹³ For more in-depth technical discussion of this the reader may consult Combettes (1993) and Combettes and Trussell (1991).

¹⁴ Present day philosophers and mathematicians erroneously use paradox and antinomy as synonyms. A paradox is an internally self-contradictory theory or model such that an accepted set of premises P plus at least one additional premise Q is such that *both* $(P \ \& \ Q)$ and $(P \ \& \ \sim Q)$ are self-contradictory. An antinomy is a pair of specious proofs of both a thesis (T) and an antithesis (A) arrived at by "proving" one by means of "proving" the other false, i.e., $(A \ \text{because} \ \sim T \ \& \ T \ \text{because} \ \sim A)$.

non-stationary system in the global sense. There are in general two ways in which this situation can arise. The first is when the system undergoes some *variation* in the parameters determining its model but does not change in the fundamental structure of the system. In effect, the result Ξ_n has *drifted* outside the solution set $\Xi^{(n-1)}$, violating a premise that all solution sets must be consistent because the system itself is unvarying. This case is identifiable because if the possible solution set "universe" is "re-inflated" (i.e., if the old results are discarded) continuation of the series of trials from that point produces a new stable solution set $\Xi^{(m>n)}$. System theorists, e.g. Rao and Huang (1993), often call this a "rescue procedure" or a "tracking procedure." The second case is where the natural system itself has undergone a change in its own internal structure such that the *a priori* knowledge of the system's structure is no longer objectively valid even though at some previous point this knowledge was correct and objectively valid. The system is said to have undergone either *variation* or *speciation*. All open-systems, e.g. living biological systems, are potentially subject to occurrences of this second kind.

V. Scientific Reduction and Model Order Reduction

One of the most effective practical tactics evolved in the practice of science over the centuries has been the division of scientific topics into a hierarchy of levels beginning with observable macroscopic phenomena and continuing downward (or, as some say, inward) toward phenomena that the scientist thinks are "more fundamental" or "more essential" according to his concept of what it means for one object to be in some sense more fundamental or more essential to understanding nature. This view is an ancient one dating back to at least Plato, although Aristotle was the first to write about it with rigor:

When the objects of an inquiry, in any department, have principles, causes, or elements, it is through acquaintance with these that knowledge is attained. For we do not think that we know a thing until we are acquainted with its primary causes or first principles and have carried our analysis as far as its elements. [Aristotle, *Physics*, 184^a10-15]

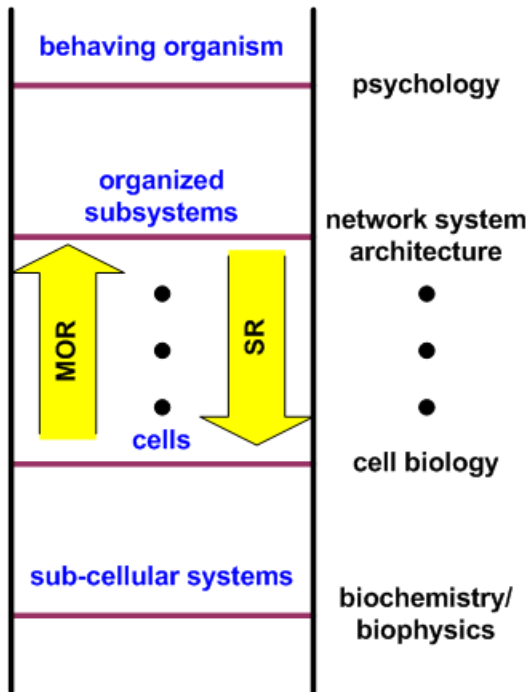


Figure 6: Ladder structure of scientific reduction (SR) and model order reduction (MOR)

This practice of migrating scientific study and theory from the level of phenomena more directly observable by our senses to levels of increasingly refined scientific constructs (e.g., atoms) is called *scientific reduction*. Historically and traditionally scientists have tended to interpret Aristotle's dictum to mean that understanding nature is synonymous with understanding things at the level of some postulated "absolute essence." The closer to some ontological "primitive essence" scientists think their science has brought them, the more "pure" or "true" science is taken to have become.

This presupposition is fundamentally flawed in three very important ways. First, its ontology-centered presupposition of some thing-like prime essence or "most fundamental entity" is an illusory idea lying far beyond the horizon of possible experience. It utterly lacks any objectively valid grounds whatsoever and is the product of mere rational induction. It belongs, in other words, to Slepian's facet B and not to the facet A of real experience. Critical metaphysics teaches us that it

is and forever will remain a mere secondary quantity of pure mathematics. Second, the practice of scientific reduction is one of abstraction *essentially*. The successes achieved through scientific reduction are purchased at the price of throwing away phenomenal knowledge pertaining to grosser objects, yet these are the objects of the *practical purpose* for the science in the first place. It is incorrect to interpret Aristotle's words in the classical tradition because in Greek physics (φυσική) *means* "the study of nature." The "principles and causes" to which he referred *include* those "principles and causes" that connect human understanding of, e.g., "atoms" with human understanding of, e.g., the design and construction of bridges or buildings. We cannot truthfully say we are studying nature if all of our efforts are made using dogmatic reductionism.

However, and thirdly, to actually be able to perform calculations and obtain answers, the number of equations must be small. A civil engineer does not, nor could he, design a bridge by performing calculations based on the theory of atoms. The objects of his science are very remote from the objects of an atomic physicist. Now, if *no one* practiced reductionism, there would be no scientific achievements. It is therefore reasonable, proper, and practical that the majority of scientists carry on their work within different relatively narrow restricted topics, and doing so is reductionism. The scientific tradition here likewise dates back to the classical Greeks. Otherwise, as someone once sagely observed, "It is difficult to study nature because there is so much of it!"

The practice of scientific reduction leads to the development of specialized *disciplines* (fields of study). Metaphorically, these disciplines can be regarded as rungs in a ladder of science-structure as illustrated by figure 6. However, if *every* scientist were a specialist working on his own particular rung, science *overall* would collapse because its various rungs cannot levitate in thin air all by themselves. But the dogma of rigid scientific reductionism promotes the evolution of isolated silos of knowledge – the special disciplines – and discourages the development of a branch of science devoted not to rungs but, rather, to the *rails* of the ladder. The task of a scientist engaged in "*rail theory*" rather than a specialist's "*rung theory*" is nothing else than the integration of science in general. He is a *generalist* rather than a *specialist*. If we liken scientific reduction to climbing down a ladder, his task involves the development of theory for this technique. But it likewise involves the development of theory for the technique of climbing *up* the ladder, i.e., finding and developing the scientific methods for re-integrating the findings on the lower rungs in the next higher rung. This task is a synthesis coordinated with analysis, a type of representation mental physics calls *anasynthesis*. Inasmuch as the practical problem of calculating useful results from this modeling (theoretical) effort must be solved for "ladder climbing" to be accomplished, his task includes what is called *model order reduction*, the science of reducing lower-level models collectively involving practically unsolvable numbers of equations to practically computable higher-level models. This, too, is illustrated in figure 6.

It is erroneous to regard model order reduction (MOR) as a form of so-called approximation theory because MOR is not a method for approximating anything. Rather, it is a methodology for learned abstraction from specific cases to produce the higher concepts that understand lower ones. If one insists on regarding MOR as approximation theory, then logical consistency demands nothing less than that *scientific reduction (SR) also be regarded as approximation theory* because the foundation of SR is likewise based on abstraction – in its case by discarding phenomena of empirical experience. If abstraction is held-to-make MOR "mere approximation," then it likewise does the same thing to SR and all pretense that SR is somehow more fundamental or more essential to understanding nature collapses under the weight of self-contradiction. Weinberg expressed the task of the generalist thusly:

The generalist, then, has certain categories of thought that, because of their general nature, are not going to fail him completely in the study of any new field. He has special words in his vocabulary, words such as stability, behavior, state space, structure, regulation, noise, and adaptation, which he can relate to the words of the specialist. . . .

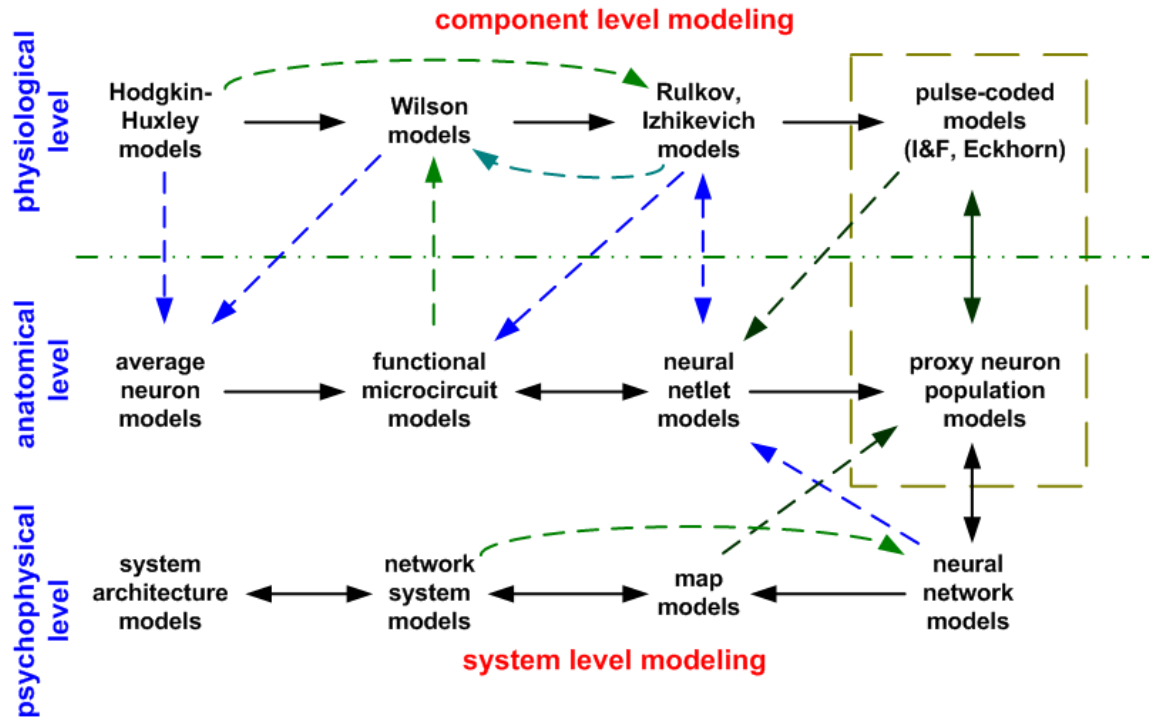


Figure 7: An alternative view of ladder structure applied to neuroscience

When the generalist encounters laws in the special field, he will often be able to relate them to the general systems "laws" he knows. He identifies the special assumptions that have made his general systems laws into laws of economics, or whatever. . .

The general systems approach, then, can engender a parsimony of thought for the study of subjects. A similar economy is introduced in the study of situations, or *special systems*. [Weinberg, *An Introduction to General Systems Thinking*, pp. 45-46]

When we narrow the general systems focus a bit and apply it to neuroscience and brain theory, we find that the methodology of the general systems thinker is expressed by better formalizing the practical description of the neuroscience "ladder" in terms of sub-system classes defined by the scope of phenomena to which each class can be directly related by experiment or observation. The variables (mathematical objects) in each class are aimed at producing Slepian's overlap with facet A on the scale of those particular kinds of experience. Figure 7 illustrates one schematic description of this ladder concept for application of general systems methodology to a science of mind-brain. Wells calls this diagram a "doctrine of a systems roadmap."

At the outset of the development of a somatic science of brain-object, we find ourselves stationed at the point in the roadmap of figure 7 labeled "system architecture models." This is because the special object of our science-to-be is the whole human being, the object-in-experience who exhibits both the phenomenon of mind and the phenomenon of body. Critical epistemology tells us that the so-called mind-body division has objective validity *only* as a mere logical division – a convenient form of categorization the scientist uses to distinguish between the sensible objects laid to "body" from the supersensible objects, e.g. of psychology, laid to "mind." The division is mathematical, belongs entirely to facet B, and must be treated as such. There is no *ontological* significance in the mind-brain division whatsoever.

The mathematical theory of brain-object is organized in a ladder structure of objective scopes for the same reason the SR/MOR structure is employed by every other science. In its case, what is

"essential" for the *science* is that the holism of its central object, the entire human being, not be lost in the process of mathematical division and scientific reduction. A Critical doctrine of method therefore requires the work to begin at the point closest to phenomenal experience with what in figure 7 is called the psychophysical level of study. This is the "rung" of system architecture models. But these models are themselves assembled from models at the next lower level – the network system modeling level. Historical practices developed over the past sixty years have come to call this level by the name "neural network theory," although strict attention and adherence to a ladder methodology would properly assign this name to an even lower rung (as suggested in figure 7).

It is at the conjunction of the systems architecture-network systems-map model rungs of this ladder where we find an enlightening and key relationship between set membership theory and an important facet of classical artificial neural network theory.

VI. Set Membership Technique, Artificial Neural Networks, and Vigilance

Present day research at all levels above that of the average neuron models in figure 7 is commonly called "neural network theory," although a decent respect for the importance of precise technical language in science properly demands it be called "artificial neural network theory" (ANNT). This is both because the models employed are, in regard to facet A, *very* artificial and because its arena of activities is properly characterized as consisting of preponderantly Platonic speculations and prejudices. It would be proper doctrine of method at this point to state the concise unifying idea under which these activities make up the practice of a science, but this cannot be done because ANNT *has no* agreed-to idea of unity fit to establish anything more than what Kant called an historical doctrine of nature. ANNT can be called a pre-science, but it is not a single, unified natural science in any sense of that term.

What do artificial neural network (ANN) models do, i.e., why do people construct them? Any interested yet dispassionate survey of the corpus of existing literature in this arena¹⁵ must conclude that its eminent researchers offer nothing more than vague descriptions of this. Malsburg and Schneider tell us,

The act of perception, in higher animals and in man, may be divided into three highly interdependent processes: segmentation, pattern recognition and integration of patterns into a scene. Segmentation separates the field of sensory information into pieces which form patterns. [Malsburg & Schneider (1986)]

Carpenter writes,

Neural network analysis exists on many different levels. At the highest level we study theories, architectures, hierarchies for big problems such as early vision, speech, arm movement, reinforcement, cognition. Each architecture is typically constructed from pieces, or *modules*, designed to solve parts of a bigger problem. These pieces might be used, for example, to associate pairs of patterns with one another or to sort a class of patterns into various categories. . . In this review I will focus on the middle level, on some of the fundamental neural network modules that carry out associative memory, pattern recognition, and category learning. [Carpenter (1989)]

As a final example, Anderson tells us,

¹⁵ I call it a "research arena" rather than a "research field" because overall it has neither discipline, canon, nor architectonic and so is not well enough organized to merit being called a "field of science." *Collectively* it is a loosely knit enterprise of applied mathematical gropings with pretensions of being a science. This is *not* to say there are no important findings arising from specific researches by specific individuals; there are. It *is* to say that *overall* it is not yet a science.

The operation of *association* involves the linkage of information with other information. Although the basic idea is simple, association gives rise to a particular form of computation, powerful and idiosyncratic. The mechanisms and implications of association have a long history in psychology and philosophy. Association is also the most natural form of neural network computation. This article will discuss association as realized in neural networks as well as association in the more traditional senses.

Neural networks are often justified as abstractions of the architecture of the nervous system. They are composed of a number of computing units, roughly modeled as neurons, joined together by connections that are roughly modeled on the synapses connecting real neurons together. The basic computational entity in a neural network is related to the pattern of activity shown by the units in a group of many units.

Because of the use of activity patterns – mathematized as state vectors – as computational primitives, the most common neural network architectures are pattern transformers which take an input pattern and transform it into an output pattern . . . In a very general sense, therefore, neural networks are frequently designed as *pattern associators*, which link an input pattern with the "correct" output pattern. [Anderson (2003)]

Elsewhere Anderson (1983) points out that the "association" he describes here constitutes the formation of mathematical equivalence classes represented by what are generally called *prototype vectors*.

I will critique this sorry state of ANNT and set out the proper Critical treatment of these issues in a later paper on the doctrine of representation. For the present purposes of this paper, it suffices to say that nominally-defined ideas such as "segmentation," "pattern recognition," "scene integration," "category learning," "association" and "associative memory," as performed by such artificial neural networks, all involve at some point the action of subsuming some *set* of vectors under a prototype vector. Mathematically, *this is functionally equivalent to forming a set membership solution set*. "Pattern recognition" is functionally equivalent to identifying which one of a divers collection of solution sets a particular pattern "belongs to."

When an ANN segments an input space of vectors by assigning them to specific prototype vectors, this is called *partitioning the input space* and is a fundamental operation performed by so-called "learning algorithms" employed in ANNs. The basic operation here is functionally equivalent to an organized aggregate of set membership operations that perform what system theorists generically call "system identification and parameter estimation" tasks¹⁶.

To put it briefly, every task present day ANNT undertakes to study *can be subsumed under the general doctrine of set membership theory*. In every ANN system that arises above the triviality of some toy problem, therefore, we find: (1) something built into the functioning of that system that corresponds to Slepian empirical uncertainty, ϵ ; and (2) some function or set of functions that constitutes the mechanism for what Combettes calls the cut set decision criteria in set membership estimation. None too surprisingly, these mathematical entities are called by a variety of names by different workers with no particular effort exerted to standardize the vocabulary of ANNT. One word, more or less synonymous with the practical meaning of Slepian's ϵ factor and introduced by Grossberg and his associates, is *vigilance parameter*. The vigilance parameter is part of adaptive resonance theory (ART), a theory developed by Grossberg and first introduced in the mid-1970s which commands our attention because of its purposive and more or less well-disciplined context in psychophysical theory¹⁷. ART itself grew out of Grossberg's work in the 1960s and early 1970s in embedding field theory (Grossberg, 1971).

¹⁶ for a fuller technical explanation of these terms, refer to Combettes & Trussell (1991) and McCarthy and Wells (1997).

¹⁷ (Grossberg, 1999)

The implication this holds is an important one. Howsoever flawed one might regard ANNT at present, its historical practices are not so far removed from a proper Critical treatment of its over-vague nominal topic. There is, consequently, sufficient reason to be optimistic that subsuming its practices (although not its flawed ontology) under the Critical doctrine of method will yield many great benefits for neuroscience and, particularly, brain theory.

And this, in a somewhat roundabout way, brings us back to the point at which this paper began: the problem of comparison and its *essential* subjectivity. To perform any of the mathematical tasks currently performed by ANNs *with objective validity in the context of brain theory* it is necessary to bring the idea of Slepian's ϵ parameter under the Critical requirements and limitations imposed on the system by mental physics. The functional congruence of ϵ with, e.g., the ART vigilance parameter points us in the proper direction. Because the basis of all comparison is grounded in the *Verstandes-Actus* of the synthesis in sensibility, and because the decision act in making every comparison rests with the process of reflective judgment in *nous*, the correct Critical conclusion is revealed immediately: *the vigilance parameter function Critically belongs to the affective subsystem of brain-object as this system stands in thorough-going reciprocity with affective judgmentation in nous.*

There is a direct *nexus* here with proper Critical architectonic in the doctrine of method. Indeed, ANNT already uses a term that describes this architectonic (albeit in other terms that are only nominally described and suffer from numerous ontological flaws). The term is *the actor-critic model* of ANN systems. Actor-critic theory in a primitive form was introduced, under another name, in 1973 and since that time has gradually come to be recognized as occupying a crucial role for the future of ANNT. Werbos wrote,

The title of this chapter may seem a bit provocative, but it describes rather precisely what my goals are here: to describe how certain control designs . . . could someday reproduce the key capabilities of biological brains – the ability to learn in real time, the ability to cope with noise, the ability to control many actuators in parallel, and the ability to "plan" over time in a complex way. These are ambitious goals, but the brain itself is an existent proof that they are possible. . .

Chapter 3 has already shown that the neurocontrol community has developed two general families of designs capable of planning or optimization to some degree over time . . . Of these two, only adaptive critics show real promise of achieving the *combination* of capabilities mentioned in the previous paragraph. [Werbos (1992)]

What must be added to this idea of the actor-critic architecture as it is developing in ANNT is the explicit recognition that, beyond its obvious *nexus* of Relation in doctrine of method, the actor-critic idea also shares the metaphysical *nexus*, i.e. the involvement of *history* as part of the doctrine of method. History in this context, and that of an ANN system, would have to be called *personal* history, by which I mean the experience acquired by an Organized Being and the genesis of this experience during its lifetime. In more familiar terminology, doctrinal method of history is to be seen in context with the affectivity capacities of the Organized Being and, for that reason, in connection with what some have called the development of "emotional intelligence." For Kant's missing history-doctrine of method, mental physics tells us to seek out this doctrine from the motivational dynamic of judgmentation in *nous* (Wells, 2009).

VII. Summary

Let me now bring this all together in overview. A science proper must have and answer to a Critical doctrine of the method in which we find four distinct headings: discipline, canon, architectonic, and history. I have not much to add here to what Kant has provided in *Critique of Pure Reason* (B: 740-883) other than specific context and relationship to brain theory.

1. Discipline. Discipline in doctrine of method has, in regard to the development of science, three synthetic *momenta*: mathematics; hypothesis; proof. One must always draw a clear distinction between metaphysical knowledge and mathematical knowledge. Metaphysical knowledge is rational knowledge from ideas and for any special science comprises the doctrine of its *applied metaphysic*. The applied metaphysic is the bridge between the empirical science itself and its grounding in Critical metaphysics proper. It is the basis for grounding the *real* understanding of the objects of the science. Its full development is, all by itself, an undertaking requiring much work and ought to be carried out by philosophers well educated in the Critical philosophy and mental physics. Metaphysical knowledge provides the grounds for all subsequent mathematical constructions.

Mathematical knowledge, in contrast, is knowledge through the construction of Objects. Without exception, these Objects are *noumena* and within the mathematical universe this construction builds we must always make a clear and distinct separation between those Objects that constitute our principal quantities and those that are only secondary quantities. Secondary quantities are not bound by the laws of physical nature but, rather, by laws of mathematics. For proper mathematics, these laws must themselves be formulated in congruence with the principles of mental physics. Mathematical objects do not lie in the plane of physical Nature, but for their ideas to ultimately mean anything the rules of their construction must be such that they can be brought to principal quantities at the horizon of possible experience, as figures 1-3 illustrated earlier. Kant noted,

Mathematics as synthetic *a priori* knowledge grounds its possibility on the fact that its concepts can be built up; for they have to do only with space and time, in which Objects of intuition can be given *a priori*. These, however, are *quanta*, thus mathematics is Knowledge of *quantis*. But it also regards quantity by means of numbers, by means of amounts which can be built up in time by counting. Yet this science cannot go farther than the sensible world, for only of this can intuition be given *a priori*. [Kant, *Reflexionen zur Metaphysik*, 18: 240]

By "space and time," Kant is referring to the pure intuitions of outer and inner sense in the synthesis in sensibility. The pure intuition of space is a process of topological structuring. That of time is a process of order structuring. Properly these are called *subjective* space and *subjective* time and are entirely different from objective space and objective time (both of which are mathematical objects). Pure mathematics is possible for human beings because the free play of imagination and understanding in the synthesis in sensibility produces objective perceptions (intuitions) without the need to call *immediately* upon what is givable through receptivity. The Critical laws of mathematics are none other than the Critical laws of thinking in mental physics, and the construction of mathematical Objects obeys the laws of transcendental Aesthetics. These laws are laws of Nature overall, but they are laws specifically of human Nature in its aspect as *homo noumenon*.

Mathematical discipline relies on firm definitions, and here mathematics has a pronounced advantage over all empirical sciences. This is because, as *made* objects, the objects of mathematics are-what-they-are because the mathematician *purposively* determines them to be so. He does so constructively, through reasoning schemes, and in accordance with *made* fundamental rules called *mathematical axioms*. In pure speculative mathematics, he is free to set for himself whatever axioms he chooses and it is this *arbitrium liberum* in his power of thinking that endows mathematical objects with that nominalism that is the root cause of Gödel's famous theorems. *However*, if mathematics is to be capable of more than mere transcendent speculation, ultimately its objects of secondary quantities must be linked to objects of principal quantities, and for the latter the rules are necessarily different and generally much more tightly restricted. This is because all mathematical

axioms employed in the construction of principal quantities *must be derived from Critical acroams*. For example, some of the axioms in the Zermelo-Fraenkel-Skolem system of axiomatic set theory are objectively valid, but the majority are not. Those that are objectively valid may be employed in the Critical mathematics of principal quantities, the others may not be. This is an important point I have previously explained in chapter 23 of *The Critical Philosophy and the Phenomenon of Mind*, and so I will not repeat this discussion and its proofs here.

2. Hypothesis. Mathematical hypothesis regarding principal quantities must be grounded in real experience because it is the conjunction of principal quantity and the idea of the real object of experience in facet A that forms the theoretical context Object figure 1 illustrates. Critical grounding means to establish clearly and distinctly the grounds in experience that lead to inference of the *Dasein* of the *noumenon* at the horizon of possible experience. Empirical science can go no farther than this point and cannot explain the *Existenz* of this Object. The task of understanding the *mathematical* nature of its *Existenz* is what falls to mathematics as a task, and for this the axioms of Critical mathematics in determining the principal quantity must be such that mathematical objects occupy a "plane of mathematical Reality" that is, metaphorically, orthogonal to the plane of facet A but such that the intersect of these two planes at the horizon of experience is understood *with real objective validity*. This is why the axioms of *Critical* mathematics must all be derived from the acroams of Critical metaphysics proper.

The situation is different for hypothesis in regard to secondary quantities. Secondary quantities are likewise constructed Objects of mathematics. These, however, are not bound immediately to the transcendental conditions of real experience. The axioms used in their construction are subject only to the laws of Aesthetics in the synthesis in sensibility and the laws of judgmentation in reasoning. The axioms, however, *must be axioms for the regulated employment of understanding in mathematical reasoning* and must make no *ontological* pronouncement nor be based on ontological presuppositions about physical Nature. All secondary quantities are problematical objects of pure mathematics, and so the speculative axiom system must itself be constructed with a strict accordance with Slepian's principle so far as inferences and implications for principal quantities are concerned.

3. Proofs. Discipline in method requires the clear distinction be made between metaphysical proofs, mathematical proofs, and experiential demonstration. A metaphysical proof is a proof following from application of the Critical applied metaphysic of the particular science. Such a proof provides the bridgework by which is realized the objectively valid theoretical context at the junction of the principal quantities of mathematical construction and the objectively valid concepts of real objects of facet A. In general, this kind of proof is to be applied to proving the real objective validity of the axioms of the Critical division of mathematics, which is to say they are proofs of axioms from Critical acroams.

Mathematical proofs, in contrast, are the proofs of *speculative* pure mathematics and differ very little from what the mathematician currently understands a proof to be. It begins with the definition of mathematical Objects, the statements of pertinent lemmas, the statement of the proposition to be proved, and the series of inferences from the conditions set by definitions and lemmas to the conditioned Object (which is the proposition that is to be proved).

Experiential, i.e. empirical, demonstration does not properly belong to mathematics at all except in those instances where the mathematician is still seeking the general idea and

calls upon illustration by special cases as a means for synthesizing the higher and more abstract idea that understands them. The new factor in doctrine of method for discipline in proof here is the solution set. Experiential demonstrations should be demonstrations of solution sets and not aimed at singular "point" solutions. In this, we may indeed stand in need of refining our present idea of what constitutes "existence and uniqueness" in mathematics.

Next let us summarize the canon for Critical doctrine of method. A canon is the embodiment of *a priori* fundamental principles of the correct use of a sure overall faculty of knowledge. There are in general four factors required for a proper canon in doctrine. First, the fundamental principles of which it is the embodiment can originate nowhere else than from the Critical applied metaphysic of the special science. Failure to meet this condition cuts mathematics loose from physical Nature and sets it adrift in what physicist-philosopher Henry Margenau once called an "island universe."

Second, the canon must be an embodiment employing a set membership theory methodology and doing so according to Slepian's principle. Third, the embodied principles must strictly maintain the *nexus* with the purpose of the science itself. At the root, *all sciences are practical*. The unity of a science is the unity of a system, and this unity is what scientific *noumena* at the horizon of possible experience are to provide for it. It is by maintenance of *nexus* with the Object-of-purpose for the special science that a clear and distinct *Realerklärung* of the Object of the science is made. The root meanings of all Objects are *practical*, not speculative or ontological.

Finally, the canon must embody clear divisions between: (1) objects of opinion (i.e., Objects held-to-be-true in understanding but without an objectively sufficient reason in experience for this holding-to-be-true); (2) objects of knowledge (i.e., Objects held-to-be-true on the ground of an objectively sufficient reason for this holding-to-be-true); and (3) Objects of experience. Objects of experience are either objects of perception or objects of judgment. The former are objects with immediate linkage to real sensuous experience through receptivity. They are adjudicated not by the process of determining judgment but, rather, by the process of reflective judgment. Therefore the concepts obtained immediately from intuition in sensibility are true in the context that they arise from the principle of Axioms of Intuition in Critical metaphysics proper and in accord with the principle of formal expedience in Nature. But for these Objects the holding-to-be-true is based upon only a subjectively sufficient reason. In the case of objects of judgment, these are the constructed Objects of thinking under the regulation by ratio-expression of the process of pure Reason. Their concepts understand the concepts subsumed under them through synthesis *a parte priori* in reasoning. The canon must understand *the ground of the judgment of the Object*.

Next we turn to architectonic in the doctrine of method. Architectonic is the art of systems, and so let us ask: what is a system? Kant provides the *Realerklärung* for this:

I understand by a system . . . the unity of manifold knowledge under one Idea¹⁸. This is the rational knowledge of the form of a whole, insofar as through this the scope of the manifold as well as the place of the parts with respect to one another is determined *a priori*. The scientific idea-of-Reason¹⁹ thus contains the purpose, to which all parts and in the idea of which they are related to each other, allows the absence of any part to be noticed in our cognizance of the rest, and there can be no contingent addition or undetermined magnitude of perfection that does not have its boundaries determined *a priori*. The whole is therefore articulated (*articulatio*) and not heaped together (*coacervatio*); it can, to be sure, grow internally . . . but not externally . . . like an animal body, whose growth does not add a limb but rather makes each limb stronger and fitter [Kant, *Kritik der reinen Vernunft*, B:860-

¹⁸ in German, *Idee*. An Idea is a regulative principle of pure Reason.

¹⁹ *scientifiche Vernunftbegriff*.

861]

The architecture of the system produced through application of architectonic in doctrine of method is a *schema* that constitutes what system theorists generally call the system *model*. Of the idea of a schema Kant tells us,

For its execution the Idea needs a *schema*, i.e., an essential manifoldness and order of the parts determined *a priori* from the principle of the purpose [of the science]. A schema that is not presented in accordance with an Idea, i.e. from the chief purpose of reason, but empirically, in accordance with contingent aims . . . yields *technical* unity; but that which arises only as a result of an Idea (where reason provides the purposes *a priori* and does not await them empirically) grounds *architectonic* unity. What we call science . . . arises architectonically for the sake of its affinity and its derivation from a single supreme and inner purpose, which first makes possible the whole. [*ibid.*, B:861]

To say that a special science has a schema is as much as to say it has what science historian Thomas Kuhn called a *paradigm*. The Critical difference is that a proper special science has its paradigm founded upon epistemological first principles rather than upon ontological prejudices.

Finally we come to history in the doctrine of method. Historical experience (knowledge through the experience of others) provides a basis for judging ideas as to their place and standing in the system. It speaks not at all to the object in regard to the object's *Existenz*, but rather to the judgment of scientific judgments. Insofar as it provides a basis for the placement of concepts in the architecture of the system, judgments of history speak to object, knowledge, and method as: (1) with regard to the object, whether it is a sensuous object (facet A) or an intelligible object (mathematical object); (2) with regard to the inception of knowledge, whether this knowledge is of empirical or rational origination; and (3) with regard to method, as to whether the method is being pursued systematically or contingently. The first concerns the problematical object, the second the actuality or non-being of knowledge, and the third the necessity or contingently of one's scientific constructs.

VIII. References

- Anderson, James A. (1983), "Cognitive and psychological computation with neural models," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. SMC-13, no. 5, pp. 799-815.
- Anderson, James A. (2003), "Associative networks," in *The Handbook of Brain Theory and Neural Networks*, 2nd edition, M.A. Arbib (ed.), pp. 117-122.
- Aristotle, "Physics," in *The Complete Works of Aristotle*, vol. I, Jonathon Barnes (ed.), pp. 315-446, Princeton, NJ: Princeton University Press, 1984.
- Berkeley, George (1734), *The Analyst or A Discourse Addressed to an Infidel Mathematician*, in *The Works of George Berkeley*, vol. III, Alexander Fraser (ed.), Bristol, UK: Thoemmes Press, 1998, pp. 1-60.
- Carpenter, Gail (1989), "Neural network models for pattern recognition and associative memory," *Neural Networks*, vol. 2, pp. 243-257.
- Combettes, P. (1993), "The foundations of set theoretic estimation," *Proceedings of the IEEE*, vol. 81, no. 2, pp. 182-208.
- Combettes, P. and H.J. Trussell (1991), "The use of noise properties in set theoretic estimation," *IEEE Transactions on Signal Processing*, vol. 39, no. 7, pp. 1630-1641.
- Feynman, Richard P., Robert B. Leighton, and Matthew Sands (1963), *The Feynman Lectures on Physics*, vol. I, Reading: MA, Addison-Wesley.

- Fogel, E. and Y.-F. Huang (1982), "On the value of information in system identification – Bounded noise case," *Automatica*, vol. 18, no. 2, pp. 229-238.
- Gödel, Kurt (1931), *On Formally Undecidable Propositions of Principia Mathematica and Related Systems*, NY: Dover Publications, 1992.
- Grossberg, Stephen (1971), "Embedding fields: Underlying philosophy, mathematics, and applications to psychology, physiology, and anatomy," *Journal of Cybernetics*, vol. 1, pp. 28-50.
- Grossberg, Stephen (1999), "The link between brain learning, attention, and consciousness," *Consciousness and Cognition*, **8**, 1-44.
- Kant, Immanuel (1787), *Kritik der reinen Vernunft*, in *Kant's gesammelte Schriften, Band III*, Berlin: Druck und Verlag von Georg Reimer, 1911. All Kant translations in this treatise were made by the author.
- Kant, Immanuel (c. 1773-95), *Reflexionen zur Metaphysik*, in *Kant's gesammelte Schriften, Band XVIII*, Berlin: Walter de Gruyter & Co., 1928. Kant quotes in this treatise were translated by the author.
- Lakatos, Imre (1978), *Mathematics, Science and Epistemology*, Cambridge, UK: Cambridge University Press.
- Lakatos, Imre (1978), *The Methodology of Scientific Research Programmes*, Cambridge, UK: Cambridge University Press.
- Malsburg, Ch. von der and W. Schneider (1986), "A neural cocktail-party processor," *Biological Cybernetics*, **54**, 29-40.
- McCarthy, Steven G. and Richard B. Wells (1997), "Model order reduction for optimal bounding ellipsoid channel models," *IEEE Transactions on Magnetics*, vol. 33, no. 4, pp. 2552-2568.
- Nelson, David (2003), *Dictionary of Mathematics*, 3rd ed., London, Penguin Books.
- Newton, Isaac (1687), *Mathematical Principles of Natural Philosophy*, 3rd ed.
- Poincaré, Henri (1905), *Science and Hypothesis*, NY: Dover Publications, 1952.
- Rao, A. and Y.-F. Huang (1993), "Tracking characteristics of an OBE parameter estimation algorithm," *IEEE Transactions on Signal Processing*, vol. 41, March, pp. 1140-1148.
- Robinson, Abraham (1996), *Non-Standard Analysis*, revised ed., Princeton, NJ: Princeton University Press.
- Slepian, David (1976), "On bandwidth," *Proceedings of the IEEE*, vol. 64, no. 3, pp. 292-300.
- Weinberg, Gerald (1975), *An Introduction to General Systems Thinking*, NY: John-Wiley.
- Wells, Richard B. (2006), *The Critical Philosophy and the Phenomenon of Mind*. available through the author's web site home page.
- Wells, Richard B. (2009), *Principles of Mental Physics*, to be published.
- Werbos, Paul (1992), "Approximate dynamic programming for real-time control and neural modeling," in *Handbook of Intelligent Control: Neural, Fuzzy, and Adaptive Approaches*, D.A. White and D.A. Sofge (eds.), NY: Van Nostrand Reinhold, pp. 493-525.