Chapter 3 Principal Quantities and Solution Sets

§ 1. Theories and Measurements

Slepian's principal quantities, standing as they do at the horizon of possible experience and, so to speak, "in between" the "real world" of phenomena and the "mathematical world" of *noumena*, belong either to something called a "theory" or else something called a "hypothesis." A hypothesis is regarded as a guess seeking to become a theory or part of a theory. Contrary to a statement I read some time ago made by a professional scientist speaking to laypersons, a "theory" and a "fact" are *not* the same thing. In science we put forth theories to *explain* facts. It is not-inaccurate to say facts are *discovered* in Slepian's Facet A and *explained* in Facet B. A fact is a phenomenon, a theory is a *noumenon*. In discussing the purpose of theories, Gerald Holton of Harvard University wrote,

> We have argued that the task of science, as that of all thought, is to penetrate beyond the immediate and visible to the unseen, and thereby place the visible into a new, larger context. For like a distant floating iceberg whose bulk is largely hidden under the sea, only the smallest part of reality impresses itself on us directly. To help us grasp the whole picture is the supreme function of theory. On a simple level, a theory helps us to interpret the unknown in terms of the known. It is a conceptual scheme which we invent or postulate in order to explain to ourselves, and to others, observed phenomena and the relationships between them, thereby bringing together into one structure the concepts, laws, principles, hypotheses, and observations from often very widely different fields. These functions may equally well be claimed for the hypothesis itself. In truth, we need not lay down a precise dividing line, but might regard theory and hypothesis as differing in degree of generality. [Holton (1952), pg. 138][1]

It has become almost a trite aphorism to say, as many do, that the key to science is to deduce the consequences of a theory, compare what the theory says with experiment or observation to see if the theory works or not, and that if experiment or observation disagree with theory then the theory is wrong. The aphorism is more or less true but it is silent on the question of *how* to make this comparison and *how to judge* a theory's agreement or disagreement with facts. Sometimes it seems more or less obvious if a conclusion disagrees with facts; but it also happens with surprising frequency that sometimes it is not. In point of fact, a single experimental disagreement with a well established theory has, historically, never overturned that theory. Experiments can and often do refute hypotheses, but a hypothesis is only a guess that has not yet achieved the loftier status of a theory. Lakatos wrote, with much truth,

> There is no falsification [of a theory] before the emergence of a better theory. [Lakatos (1978), pg. 35]

A professional who spends a significant amount of his time working in the laboratory might, from time to time, reflect that theories he works with and measurements he makes seldom provide him exactly the same numbers. A very essential purpose of having students spend considerable time in the laboratory is to acclimate them to this state of affairs. Sometimes this instructional purpose miscarries; this is the message implicit in the old saying, "That might work in theory but not in practice." But these disagreements between theory and practice almost never cause scientists or engineers to reject a theory. To the extent that Lakatos' statement is true, why is this? Holton wrote,

---

[1] Holton's book is an introductory college science textbook. What is quoted above is taken from an entire chapter devoted to "the nature of scientific theory." In science and in engineering education today there seems to be a rush – and I would say too much of a rush – to dive into the fine technical details of science without providing the student with an adequate understanding of what science is, what theories are, and why they are the way they are. Of course, this is only my opinion on our state of education; but it is an opinion arrived at through my experiences with typical students.

> When we examine why scientists as a whole, in the historical development of science, have favored or rejected a given theory, we may discern a few criteria which seem to have implicitly or unconsciously dominated the slowly evolving process of decision. (But by no means must we suppose that a theory is ever necessarily rejected solely because it fails to answer perfectly to one or another of the questions in the listing that follows.)
>
> Three qualifications have already been cited. (1) A fruitful theory correlates many separate facts, particularly the important prior observations, in a logical, preferably easily grasped structure of thought. (2) In the course of continued use it suggests new relations and stimulates directed research. (3) The theory permits us to deduce predictions that actually check with experience by test, and it is useful for solving long-standing puzzles and quantitative problems. . . .
>
> The history of science has shown us the prosperous theories frequently have some of the following additional properties:
>
> (4) When the smoke of initial battle has lifted, the successful theory often appears to have simple and few hypotheses – to the extent that the words "simple" and "few" have any meaning at all in science. The test of a theory comes through its use; hence the survival value of economy. A theory which needs a separate mechanism for each fact it wishes to explain is nothing but an elaborate and sterile tautology.
>
> (5) Ideally, the assumptions should be plausible – to contemporary scientists, of course – even if they are not immediately subject to tests; and the whole tenor of the theory should not be in conflict with current ideas. When this could not be arranged, the theory faced an often stormy and hostile reception and had to submit to long and careful scrutiny before its general acceptance. [Holton (1952), pp. 142-143]

Holton's summary speaks directly to Lakatos' dictum. A "fruitful" theory is one that has been used many times to satisfy its users' purposes and achieve useful outcomes. The longer and more often it is used, the more these benefits are multiplied. If it subsequently were to be rejected and abandoned because of some major disagreement between its predictions and actual observation, all these prior useful benefits – and, in particular, the psychological satisfaction of thinking we know enough about nature to undertake bold adventures in engineering or science – are, psychologically at any rate, cast aside. Holton noted,

> [Great] ideas arise only very rarely, compared with the large number of workable or even great ideas conceived within the *traditional* setting, so that the individual scientist is naturally predisposed to favor the traditional type of advance which he knows and believes in from personal experience. He . . . quite rightly must defend his fundamental concept of nature against large-scale destruction, particularly at the earlier stages, where the innovators cannot present very many results and confirmations of their new ideas. [*ibid*., pg. 143]

But along with all this, there is another factor at work which cannot be lightly dismissed. Most scientific observations are made, and are only possible, through the use of measuring instruments. For all but the most simple of these, those instruments are themselves fairly sophisticated devices whose designs are themselves the product of theories. Consider, for instance, so simple an instrument as a voltmeter. Figure 1 presents an electronic schematic diagram of an "old fashioned" voltmeter that employs a permanent magnetic moving coil (PMMC) display for its readout mechanism. Its electronic circuitry can properly be regarded as *approximating particular mathematical equations* that are *theoretically* descriptive of the amount of deflection of the PMMC's needle that would be caused by particular applied voltages.

Human beings cannot see electric voltages. The voltmeter is a device for converting voltage to something visible, namely, a mechanical deflection of a needle whose amount of deflection corresponds to a force that serves as a measure of the energy associated with electric potential. This *conversion* of one physical phenomenon to another related physical phenomenon is an example of a measuring instrument *extending the range* of the horizon of possible human experience. However – and this is my key point – this extension is only possible by mathematical relationships and the meter's output is *a representation of*

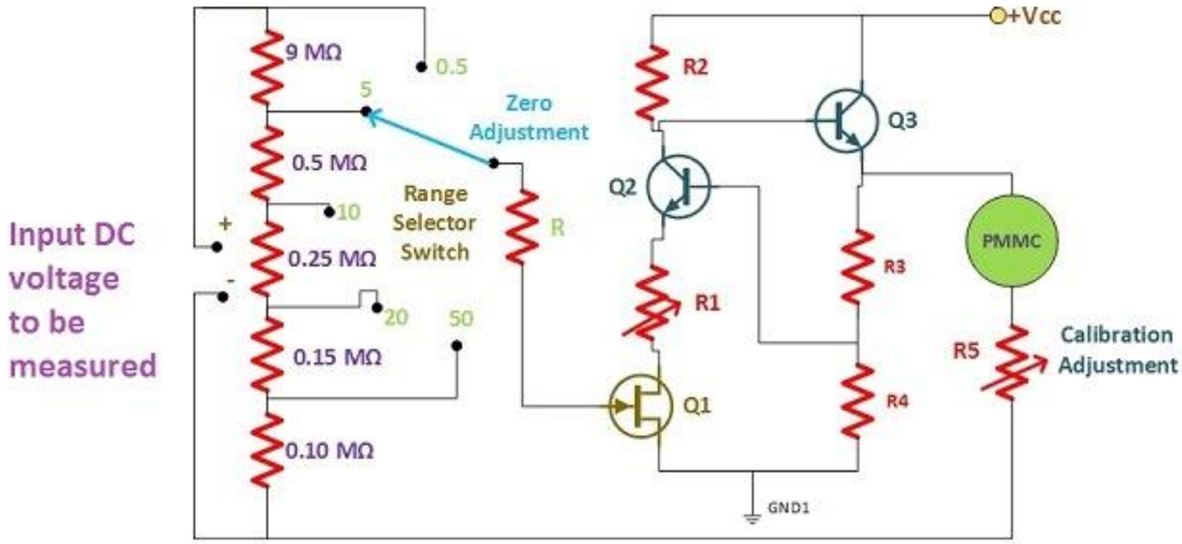

Figure 1: An "old fashioned analog" DC voltmeter employing a permanent magnet moving coil (PMMC) as its readout device. The input voltage to be measured produces a calibrated deflection of a needle in the PMMC, and the amount of this deflection is read from a scale in the PMMC calibrated to convert the amount of deflection to volts.

a noumenon. To "measure a voltage" using this instrument really means to compare an appearance of *the effect* of one *noumenon* (the meter readout) to *theoretical properties* of a different *noumenon* (the electric voltage). Therefore, what one is actually doing is making a comparison and judgment of two *different* principal quantities to one another. There is always this subtle but epistemologically significant difference between "what is observed" and "what is measured," and this difference is not confined to electronic test instruments. To appreciate this is so, consider another example provided by Holton:

> If the scientific problem were to find the length or volume of a cubical block, its electric charge, or chemical composition, the specialists called in to solve this problem would all appear to understand clearly what is wanted because they all would independently go through quite similar manual and mathematical operations to arrive at their answers, even though these answers may not completely coincide. . . . As a consequence, their findings for, say, the length of one edge of a block may read 5.01 cm, 5 cm, 4.99 cm, and 5.1 cm. The impressive thing here is not that they seem to disagree on the exact length or on the accuracy, but that they do agree on the type of answer to give. They do not, for example, say "as long as a sparrow in summer," "five times the width of my finger," "as long as it is wide," and "pretty short." Evidently the words "length of one edge of a block" mean the same things to all four experimenters, and we may be fairly sure that on consultation sooner or later all four would convince themselves of one single result, possibly 5.04 ± 0.04 cm. It should not matter which meter stick they happen to pick up, for all manufacturers of measuring equipment are under legal obligation to check their products against carefully maintained universal standards. On the whole, almost every measurement made in the laboratory is in some sense a comparison between the observables and some well-known standard accepted by the whole community. Much of the success and rapid growth of science depends on these simple truths, for it is clear that in some circumstances the energies of the investigators are not regularly drained off by fruitless arguments about definitions and rules of procedure [*ibid*., pp. 221-222].

It should not surprise you Holton's four different experimenters report four different readings. Back in chapter 1 we had an example of using a ruler to measure the length of the hypotenuse of a right triangle. The finding was that what we knew from the measurement was only an interval, i.e. "3.60 inches < length < 3.61 inches." Another way to say this is "length = 3.605 ± 0.005 inches." Pythagoras' theorem tells us "the length should be 3.6055512… inches." This *theoretical* value departs very little from the *empirical* value of 3.605 ± 0.005 inches and it lies *within the range of the measurement*. This is a simple example of what is meant by saying "a fruitful theory *correlates* many separate facts."

Not many people will find this example to be disturbing. However, the more complex the measuring instrument becomes, and the bigger the role theory plays in the design of the instrument, might well – and probably should – start to make us think a bit more about how much to trust the "data" the instrument gives us. Years ago I was working with a close colleague – a physicist – on the development of a new high technology device used in disk drives. We were discussing a puzzling experimental result one day, and my colleague voiced a concern about what he called "theory-laden data." In this particular case, the data we were looking at resulted from a very complicated measurement method and he was correct to worry a bit about *how much theory had to be correct* if our data was really meaningful. This was just one special case of an issue found in almost every laboratory instrument and almost every modern scientific measurement. It affects how we interpret and understand all correlations between mathematical principal quantities and our empirical experiences in the laboratory.

There is another aspect of all this I would like to point out. In the examples above, the fact is that our "answer" to the measurement problem is actually a number *and* a range. It is *not* just the number "3.605." Only *theories* provide as an answer just *one* single and unique number. This one single number can be called a "point solution" by an analogy similar to how a geometric "point" was described in chapter 1. It is typical to dismiss answers like 3.605 ± 0.005 as a mere reflection of the measurement accuracy of our instruments. But perhaps there is a bit more to this than "meets the eye."

Let us suppose we obtain a reading from the voltmeter in figure 1 of, say, 4.5 volts. Figure 2 illustrates a concept diagram of how a PMMC readout device is constructed. When a current runs through the movable coil, the coil rotates because there is a torque acting on it due to the permanent magnet. This rotation is opposed by fixed mechanical springs. A pointer needle attached to the rotating coil points to a scale reading that is calibrated to how much deflection a given voltage would theoretically cause. What the voltmeter does is convert the electric potential energy in the voltage being measured to mechanical potential energy. The person taking the measurement "reads" the voltage by looking at the scale position of the pointer. As you can surmise, the last step in the measurement process is reduced to the same kind of measurement as when we use a ruler to measure a distance.

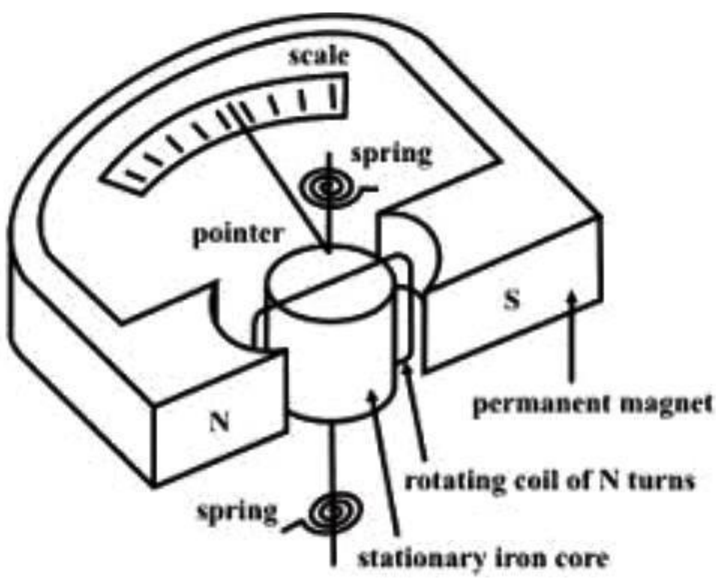

Figure 2: Conceptual illustration of how a PMMC display is constructed.

Typically, the pointer needle will look very still after it reaches its final position. However, if you were to look at the scale reading using a powerful enough optical magnifier, you would see the pointer jiggling around a little. An electrical engineer will call this jiggling "noise" and regard it as an unwanted random signal. (The word "random" is synonymous with "unpredictable"). Statisticians (mathematicians who specialize in statistics) will call 4.5 the "mean value of the pointer's position" and call the jiggles "random variables." Empirically, it appears to be the case that "noise" attends the measurement of all but the most simple operation, i.e., counting. MacDonald provides a brief history of the scientific study of noise [MacDonald (1948)]. Holton wrote,

> Lastly, you may be puzzled that a simple measurable like length should not be given with perfect exactness, that we should have said "the edge of this block is 5.04 ± 0.04 cm long." The term ± 0.04 cm is called the Probable Error in the measurement; that is to say, the chances are fifty-fifty that the next measurement of the same edge will read between 5.00 cm and 5.08 cm. Evidently all measurements, apart from simple counting, will contain some error. But although this uncertainty may be slight, how can we dare build an exact science on concepts defined by necessarily uncertain measurements? The paradox is resolved by two recognitions. First, this word "error" does not have in science the connotations *wrong*, *mistaken*, or *sinful* which exist in everyday speech. In fact, a measured value can be regarded as "exact" only if we know also just what range of values to expect on repeating the measurement. Secondly – and we must repeat this frequently – science is not after absolute certainties, but after relationships among observables. We realize that the observables can be defined and measured only with some uncertainty; and we shall not demand more than that of the relationships between them, thus leaving the search for absolute truths to other fields of thought. [Holton (1952), pp. 223-224]

If one wishes, Holton's thesis here can be given an Orwellian-sounding note by saying this is as much as to say "exact is inexact"; but to say so misses the point. Science progresses by means of theories; but *no* theory takes into account everything in the universe. For that reason, every theory is a "model" of nature of some kind (where by "model" I mean "a representation that mirrors, duplicates, imitates, or in some way illustrates a pattern of relationships observed in data or nature"). Recall from chapter 2 that Slepian said Facet B "is a mathematical model and the means for operating with the model." He drew the distinction between principal quantities and secondary quantities of Facet B by saying,

> There are certain constructs in our models (such as the first few significant digits of some numerical variable) to which we attach physical significance. That is to say, we wish them to agree quantitatively with certain measurable quantities in a real world experiment. Let us call these the *principal quantities* of Facet B. [Slepian (1976)]

In light of the discussion above, perhaps you can now better appreciate the epistemological significance of Slepian's argument. In chapter 1 a large part of the text was devoted to asking "what is a number?" You most likely noticed at the time that while much was said about "numbers" the chapter concluded without providing an answer to this question. It did, however, arrive at some key characteristics of them:

1. Numbers are outcomes of measurements;
2. Number-objects are defined by actions and therefore are *practical* objects;
3. Numbers are given their meanings by practical operations; and
4. Psychologically, Piaget found the idea of a "number" regarded as a *genus* is the outcome of synthesizing a generalization of equivalence ("all the elements of a set are equivalent to each other") *with* a generalization of seriation ("the elements of the set are different from each other and their differences allow them to be placed in some definite serial order or arrangement"). "Number" is a synthetic combination of these two ideas.

The first three of these statements seem easy enough to grasp. The fourth, on the other hand, is markedly less so. The general idea of a number is a representation of an Object that contains, on the one hand, objects that are "the same" as each other, e.g., {"one," "two," "three"} are "all the same" (they are all words). But *at the same time* they are *operationally* different in a way that imposes an *ordering* structure on them – for example, extend your index finger ("one") *and then* also extend your middle finger ("two") *and then* also extend your next finger ("three"). Smith tells us,

> In the Malay and Aztec tongues, the number names mean literally one stone, two stones, three stones, and so on; while the Niuès of the Southern Pacific use "one fruit, two fruits, three fruits," and the Javans use "one grain, two grains, three grains," all these being relics of the concrete stage of counting. [Smith (1923), pg. 8]

The words in the set aren't "numbers" *until and unless* some differentiating ordering operation is combined with them (e.g. pointing at a stone while saying "one" and then pointing at another stone while saying "two" and then pointing at yet another stone while saying "three"). As Piaget put it, the idea of "number" requires a person to think of the terms in the set as equivalent but, at the same time, *keep them distinct* from each other in an asymmetric relationship (e.g., $1 < 2 < 3$). No meaningful measurement is possible until a person synthesizes the identity of his "number terms" with an ordered plurality of these terms.

As you hopefully can appreciate at this point, theories, measurements, models, and "number" (as a genus) are more intimately interconnected than we usually think when regarded from an epistemology-centered "way of looking at the world." In Critical epistemology, all meanings are, at root, practical. Might it not be, therefore, that the answer to "what is a number?" can be found by examining this inter-connectedness and thereby understanding how and when Facet B quantities are objectively valid when we use them to better understand empirical science? Certainly we can frame this as a hypothesis and see where it takes us. Let us start with Holton's statement, that a measured value is exact only if we also know what range of values to expect on repeating the measurement, and pair it up with Piaget's finding.

§ 2. New Math

In Holton's hypothetical example of four experimenters independently measuring the edge of a cubical block, it is supposed they came up with a set of measurements {5.01 cm, 5 cm, 4.99 cm, 5.1 cm}. It is then further supposed these experimenters got together, compared their results, and agreed "the length of the cube edge is 5.04 ± 0.04 cm." He didn't offer any explanation for how the 5.1 cm result (which is larger than $5.04 + 0.04 = 5.08$ cm) or the 4.99 cm result (which is smaller than $5.04 - 0.04 = 5.00$ cm) figured into the final answer agreed upon by his four experimenters other than to say the next measurement has "a 50-50 chance of being 5.04 ± 0.04 cm." Holton had his four experimenters all agreed to *say* the *objectively valid* answer was 5.04 ± 0.04 cm. No distinction is thereafter made between their results. In effect, they (or Holton) are saying "*for all practical purposes* '5.04 ± 0.04' is a number."

Now, most of us, it is true, don't usually think something like 5.04 ± 0.04 is "a number." But it is also true that most of us most of the time *don't need* to think of it as a number. A parent doesn't tell the census taker "we have 2 ± 0.25 children." On the other hand, according to the company Statista, in 1970 the average American family had 2.28 children under the age of 18; it says in 1980 there were 1.91 children; for 1990 their figure is 1.83; for 2000 they said it is 1.87; and for 2010 they said it is 1.88 children. The company doesn't report a "±" figure for any of these, nor does it explain how it arrived at these figures. From the viewpoint of empirical science, *these figures are meaningless* because we have no idea how to use them correctly and no guidelines or rules for how to deduce consequences of any kind from them. We can look at this *collection of figures* {2.28, 1.91, 1.83, 1.87, 1.88} and call them "numbers" if we wish to because they look the same as other things we are taught to call "rational numbers"; but how were they arrived at? what do they *imply*? how can we *use* them to determine something else? According to the U.S. Census Bureau, in 2010 there were 78,833,000 households in the U.S. and they had 74,718,000 children under 18 in them [U.S. Census Bureau (2011), Tables 64, 69]. Divide the latter figure by the former and you get about 0.948 children under 18 per household. This certainly isn't the same as 1.88 so Statista's figure doesn't represent this. We can say Statista's figures are "numbers" if we choose, but ***we cannot call them principal quantities*** because we don't know how to relate them to any context-providing phenomenal Objects. We could, of course, say they are all "rational numbers" and place them in a pure mathematics[2] context, but doing that does not establish any context for them in empirical science. Principal quantities *must* have some relationship with sensible *phenomena*.

Let us expand Holton's hypothetical example a little. Suppose his four experimenters next performed independent measurements on another edge of the block, and suppose they report 5.21 cm, 5.15 cm, 5.25 cm, and 5.20 cm as their results. They might then get together, discuss their individual results, and then agree upon a final result of 5.20 ± 0.05 cm. Now we can compare the two results, "5.04 ± 0.04 cm" and "5.20 ± 0.05 cm," and note that all four results from the second experiment are greater than all four results of the first experiment. We would be justified in concluding "5.04 ± 0.04 cm" < "5.20 ± 0.05 cm." If we do so conclude, then we have *seriated* the two "funny numbers" "5.04 ± 0.04 cm" and "5.20 ± 0.05 cm." We can say not only that they are both "funny numbers" of the same kind but also they both fall under Piaget's condition for something we can call a "number in general." This is to say *at least some* "funny numbers" of the form $x \pm y$ are *species* of the genus "number." Thus, we have just invented a new "class" of numbers.

We can also draw at least one empirical conclusion. Because "5.04 ± 0.04 cm" < "5.20 ± 0.05 cm" means the two measured edges are not the same length, the block that is being measured is not a cube because the *definition* of a cube requires all the edges to be the same length. Thus we can *apply* our new class of numbers to objects of sensible experience. *This makes them principal quantities*.

Having come up with the idea of a class of "funny numbers" of the form $x \pm y$, a mathematician's job is not yet done because there is still a lot we'd like to know how to do with our funny numbers. What rules do we want in determining whether some particular $x \pm y$ is to be included in the class? Can we *add* two of these numbers together? and, if so, what is the result? and is this result also a member of the class? Can we define "division" and "multiplication" operations that can be performed on this class? and, if so, how? Assuming we can define one or more "addition" operations on them, does the sum of three "funny numbers" obey the associative property of addition? Can we work out an algebra for them? These and many other ideas of *secondary* quantities for numbers in this class still remain to be worked out.

These are examples of some of the kinds of work carried out by mathematicians in what is called "pure mathematics." It subsists in *inventing, constructing, and analyzing* what are called "formal systems of mathematics." Preparata & Yeh wrote,

> Roughly speaking, a formal system consists of notions, or concepts, and of assertions about their properties. It would be ideal to have a precise method that established the absolute nature both of the definitions of the concepts and of the validity of the assertions. However, a moment's reflection shows that such an ideal is unattainable since making use of other concepts is inherent in the nature of explanation. And in order to explain the meaning of these new concepts, one has to resort to different concepts (in order to avoid vicious circles), and so on. Hence one finds oneself in a process which apparently cannot be brought to an end. To see how common the process is, one can look up almost any word in a dictionary, then look up the words used for its explanations, and so on. It usually takes only a few steps for the original word to reappear in an explanation. . . . In constructing a formal or mathematical system, we first select a collection of concepts that seem to us immediately understandable (usually because of their closeness to immediate experience). [Preparata & Yeh (1974), pp. 2-3]

Most of the work carried out in pure mathematics centers on ideas of secondary quantities in Slepian's Facet B. For that work to be applicable in, say, engineering or physics, principal quantities have to also be

[2] "Pure" mathematics is a science but it is not a science of nature ("natural science").

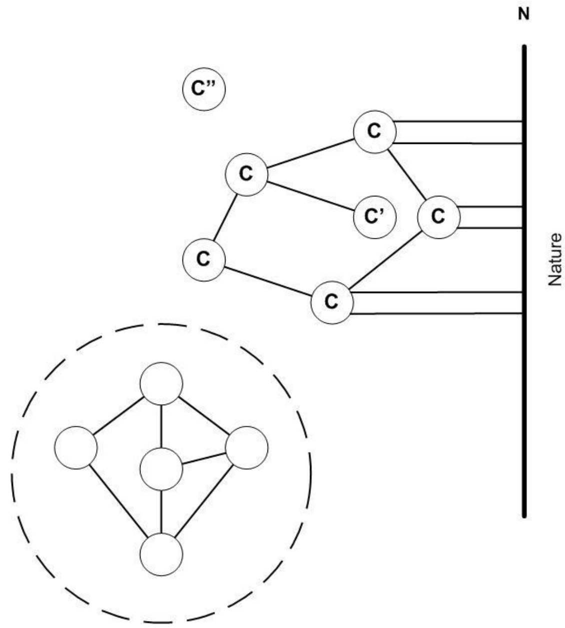

Figure 3: Illustration of Margenau's "construct theory" of nature and reality. The solid dark line N represents what he called "the plane of the immediately given" (i.e., immediate sensuous experience). The constructs C shown connected by double lines to this plane N correspond to Slepian's principal quantities of Facet B. The remaining constructs C correspond to Slepian's secondary quantities of Facet B. The unlabeled circles within the dashed circle represent constructs of pure mathematics for which there are no known principal quantities. They constitute an "island universe" of pure math constructs lacking any known connection to nature. [Margenau (1977), pp. 75-101]

defined and worked out.

Some mathematicians do not particularly care if their work has any "real world" applications and devote their time to studying mathematical structures somewhat as if they were just amusing puzzles that are stimulating, entertaining, and challenging. From time to time some of them tell their students, "You'll never have to worry about this in the real world" (and, at that point, most students of engineering stop listening). However, in the history of mathematics it happens that, sooner or later, somebody *does* invent principal quantities that connect a mathematician's work to the "real world" and – all of a sudden – what was just an exercise in pure mathematics finds itself engaged in what is usually called "applied mathematics." If the new application is sufficiently important to a sufficient number of people, the "pure mathematician's theory" ends up being taught to non-mathematicians in disciplines outside the math department. Non-Euclidean geometry is an example of this. There are actually an unlimited number of them. They arose from efforts by mathematicians to prove that Euclidean geometry was the only possible geometry. Then, to everyone's shock and surprise, Janos Bolyai (1823) and Nicholai Lobachevsky (1826) discovered that perfectly self-consistent non-Euclidean geometries could be formulated [Davis & Hersh (1981), pp. 217-223]. Bernhard Riemann later developed elliptic geometry (known as Riemannian geometry) in 1854. Non-Euclidean geometries were "island universe" constructs until Einstein used Riemannian geometry in 1915 for the theory of general relativity. Einstein's work provided the link between Riemann's geometry and nature. More recently, Benoit Mandelbrot introduced yet another geometric construct with "real world" connections, which he called fractal geometry [Mandelbrot (1977)].

Definitions of principal quantities in science always involve some practical procedure or operation by which numerical values for the principal quantity are obtained. The procedure or operation is what *makes* the connection between mathematics and the natural "real world." Holton wrote,

> If we define as "meaningless" any concept not definable by operations, it follows that not only concepts but whole statements and even intelligent-sounding, deeply disturbing questions may

> turn out to be meaningless *from the standpoint of the physical scientist*. . . . As has been implied throughout [this book], not all concepts in science can be defined by *manual* operations. Instantaneous velocity, for example, was defined by the slope of a straight line tangent to the distance *vs*. time graph; this is a mental or mathematical operation. But each such definition becomes a fairly unambiguous set of directives to all scientists who have worked in the field, and this explains why their vocabulary is not regularly the subject of dispute.
>
> You will find it revealing – more than that, almost indispensable to a true understanding of physical concepts – to make for yourself operational definitions of all important terms, such as we have made in previous chapters for acceleration, force, mass, and so on. In this task also it will be wise to keep in mind Bridgeman's simple statement: "The true meaning of a term is to be found by observing what a man does with it, not what he says about it." [Holton (1952), pp. 225-226]

I confess I am not confident students graduating from today's science or engineering curricula are leaving it with this important lesson in tow because I am not confident most of today's educators are teaching this lesson to them. I do think *failure* to teach it is educational malpractice.

Let us take a moment to recap the thesis that has been developing in this treatise. The "reality" of a secondary quantity of Facet B is determined by the polysyllogisms and disjunctive inferences that establish a series of concepts by which that quantity is connected to principal quantities of Facet B. The "reality" of a principal quantity is established by rules and operational definitions that determine its application to sensible phenomena. The result is a *model*, i.e., a system of Objects, postulates, data, and inferences presented as a mathematical description of an entity or state of affairs.

Mathematicians and other scientists have a habit of talking about what they call "primitive terms" (often abbreviated to just "primitives"). Usually they say or imply that a "primitive" is something that cannot be "defined." For example, Whitehead & Russell tell us,

> Since all definitions of terms are effected by means of other terms, every system of definitions which is not circular must start from a certain apparatus of undefined terms. It is to some extent optional what ideas we take as undefined in mathematics . . . Hence we can only say such and such ideas are undefined in such and such a system, not that they are undefinable. Following Peano, we shall call undefined ideas and the undemonstrated propositions *primitive* ideas and *primitive* propositions respectively. The primitive ideas are *explained* by means of descriptions intended to point out to the reader what is meant; but the explanations do not constitute definitions because they really involve the ideas they explain. [Whitehead & Russell (1910), vol. 1, pg. 95]

But "undefinable" is not the same thing as "unexplainable." A *practical* procedure or operation informs us of the *real meaning* of a term (like "number") and explains to us, in effect, what the "nature of a primitive term" is.

In Critical epistemology, a primitive term is a technical term that cannot be reduced to and explained by more fundamental terms and has only a practical *Realdefinition* or *Realerklärung*. A *Realerklärung* ("real explanation") is an explanation in terms of those powers and processes of the phenomenon of mind by which an understanding of nature is constructed, structured, and understood, and which makes the objective reality of the concept distinct [Wells (2006)]. A *Realdefinition* ("real definition") is a practical definition that contains a clear mark by which the object can always be recognized and makes the concept to be explained *usable in application* [*ibid*.]. Both these terms come directly from Kant [Kant (1787) B: 755-766; Kant (1800) 9: 140-145]. He wrote,

> As the expression itself imparts, *define* must properly mean simply to exhibit originally the exhaustive content of a thing within its boundaries. Given such a requirement, an *empirical* concept cannot be defined at all but only *explicated*. For since we have in it only some marks of a certain variety of objects of sense, it is never sure whether under the word that designates the same object one does not sometimes think more of these marks but another time fewer of them. [Kant

(1787) B: 755]

Mathematical objects can be defined (operationally); *phenomenal* objects cannot be defined and can only be explained or described (and this non-exhaustively) on the basis of experience because we do not know if or when some *new* experience will overturn a definition of a phenomenon.

"Number" as a genus is often treated in mathematics as such a primitive term and, frequently, in such a way as to suggest to the reader that the writer's viewpoint is ontology-centered. From such a metaphysic "number" truly is undefinable and unexplainable. Hence, for example, the ancient Pythagoreans came to hold a mystical and semi-religious regard for numbers. Some people today hold what I would have to call a semi-religious regard for "numbers" and express shock if you "mess with the number system" (as if there were only *one* of them).

However, from an epistemology-centered view, the genus has a practical *Realerklärung* – specifically that presented in the four bullet points on page 48. These speak to the psychology of "where numbers come from" and how they are practically applied to the sensible world of phenomena.

Mathematics deals with mathematical Objects, and these Objects are, one and all, products of the human power of imagination. Unlike empirical concepts, they are defined, either in terms of one another if they are secondary quantities or practically by operations, such as our "funny numbers of the form $x \pm y$", if they are principal quantities. When new Objects of mathematics are defined, one might be confronted with a need to develop *new* mathematical structures ("formal systems") to advance their applicability for science. It is not uncommon for scientists and engineers – when confronting something new or different from what has been dealt with before – to try to treat the problem or question by *forcing the problem* to fit the "old math" the scientist or engineer already knows and is comfortable with using. But, once upon a time, "old math" *was new* and was created in response to some real world problem or interest. When a scientist or an engineer tries to stuff a new problem forcibly into an old math structure, then he is not actually working on the problem he started out to solve but is instead working on an entirely different problem and is behaving as if all the math that could exist already exists – and that is certainly not true.

Forcing a science or engineering problem to fit the math one has already at hand is, I submit, a mistake. It is the phenomenal world the scientist or engineer is concerned with. When math *that already fits* the problem exists, using that math is perfectly correct. But *distorting* a problem to fit a particular formal system of mathematics turns an empirical problem into a mathematics problem disconnected with the phenomena the scientist or engineer actually wants answers for. To the consternation of many students, mathematical formulas do not come with a user's manual to tell them "use me here and here but not over there." One frequent result of this is that a student will just pluck an equation from his textbook and apply it to his assigned homework problem *even when it does not apply*. The crucial skill he lacks is the skill of *figuring out what math applies to the problem he is working on*. And he lacks it because he is not being taught that *this* is a skill needed for him to *become* a scientist or engineer.

It is just simply the case that *sometimes we must make mathematics to fit the problem*. When you do so, you can be said to be constructing and using "new math." It has happened many times in the history of science and engineering: Newton's calculus; Fourier's transforms; Heaviside's operational calculus; Shannon's "entropy algebra"; Zadeh's "fuzzy logic"; Mandelbrot's "fractals"; and other examples. There is no reason to suppose this won't continue to happen. Some people might find this prospect dreadful; personally, I have always found it to be a great deal of fun. Poincaré wrote,

> It may seem surprising that sensibility should be introduced in connexion with mathematical demonstrations which, it would seem, can only interest the intellect. But not if we bear in mind the feeling of mathematical beauty, of the harmony of numbers and forms and of geometric eloquence. It is a real aesthetic feeling that all true mathematicians recognize, and this is truly sensibility. [Poincaré (1914), pg. 59]

I agree with him.

§3. Solution Sets and the Set Membership Paradigm

What does uncertainty and "noise" in empirical measurements imply for mathematics? Does it *require* scientists to resort to some "new math" to adequately solve problems or undertake research in science? The answer to this question depends on the circumstances and whether or not "old math" yields results that are satisfactory, i.e., that meet the needs and purposes of the scientist or engineer.

The sorts of problems students encounter in school clearly don't seem to require it, and education in science or engineering seems to get along quite well with classic "old math" mathematics. On the other hand, the school setting is a carefully controlled environment and almost all of a student's experience in the laboratory is designed to serve purposes of pedagogy and training. Usually about the only place where students are *compelled* to deal with uncertainty in data or observations is in a course on statistics or, perhaps, in some special introductory course on laboratory equipment and practices. Here students are made aware of the real existence of the phenomenon of uncertainty but, in most educational institutions, instruction is such that most students come to regard uncertainty and noise as annoying inconveniences unrelated to the theory they are learning, and instruction tends to ignore puzzles such as the bandwidth paradox and any implications such paradoxes can hold for theory and research. Moreover, in the past few decades *simulations* have begun to replace *laboratory* instruction in many institutions of education. But simulation is not experiment. Simulations are *entirely* mathematical exercises and do not expose the student to the characteristics of actual nature. Simulations reveal the consequences of *theory*; experiments reveal the face of *nature*.

At the same time, the mathematics instruction most students of natural science and engineering receive at the undergraduate level does not even hint that mathematics systems are invented and designed by mathematicians. The pedagogy tends to reinforce the Euclidean myth, unintentionally or not, and it tends to coronate an illusion: that answers to scientific problems are always singular – i.e., that they have one and *only* one solution. But, as we saw in Slepian's answer to the bandwidth paradox, this is not actually true in general. There are phenomena and circumstances in which multiple "correct answers" exist – all of which are consistent with all the measured data at hand and everything that is known (or believed to be known) *a priori* about the system. When this is so, any one of these answers is as good (correct) as any other. The collection of such answers is called a "solution set" because all of them "solve the problem" insofar as the limited horizon of human experience permits. It is epistemologically meaningless to ask "which one of the answers in the solution set is *really* the correct one?"

Uncertainty in measurements and observations of physical phenomena (e.g. jiggling scale readings of the PMMC voltmeter) oftentimes is only a minor problem in scientific or engineering undertakings. While it can raise some philosophical issues, much of the time these issues do not prevent the scientist or engineer from satisfying his purposes and "solving the problem at hand." The jiggles in the readings are dismissed as "noise" or, if they are too large to be safely dismissed, observations are broken down into two parts: one part pertaining to hypotheses concerning the "Object being measured" and the other part pertaining to "the sources of noise" and to means of reducing the uncertainty "introduced by noise." This time-honored approach works well enough in these cases from a practical point of view.

There are many problems in science and engineering that fall into this class. Observers tend to habitually adopt maxims of thinking about them that can be called "Platonic" because observers' unexamined presuppositions about the system and "the nature of noise" fall in line with Plato's *noumenal* ontology-centered metaphysic. Plato symbolically expressed his metaphysic in his famous "myth of the cave" [Plato (date uncertain), Bk. VII]. In this myth, we are likened to men who have lived all their lives inside a cave and have never seen anything except shadows on the wall cast by things as they pass by outside the mouth of the cave. The men mistake these shadows for reality:

> *Things*, Plato says, with an expressive metaphor, are *shadows of the Ideas*. Shadows are signs of things and they can make one aware of the existence of things. . . . In the seventh book of the *Republic*, Plato relates a myth of astonishing power, in which he represents symbolically the situation of man in relation to philosophy, and, at the same time, the structure of reality. . . . The cave is the world perceived by the senses, and its shadows are the things of the world of the senses. The outside world is the true world, the world perceived by the mind, or the world of the Ideas. [Marías (1967), pp. 48-49]

For Plato, the seemingly greater permanence of the "Object" and seemingly greater changeability of the "noise" would be a sufficient ground for dialectically making a logical distinction between them and calling the Object "one thing" and the noise "another thing." Aristotle, on the other hand, would tell us we have no empirical basis for dividing up the observation this way, and if we do so this can only be a guess because "Object-and-noise" are always seen together and are to be regarded as *one* thing whose accidents of appearance *only* are perceived as changing. Both of these "ways of looking at the world" are ontology-centered; both, in a manner of speaking, "jump to ontological conclusions"; and both, ultimately, resort to some theology for their final "explanations" [see Marías (1967), pp. 43-53; 62-74].

Modern science and engineering, on the other hand, have learned to eschew reliance upon what Newton called "occult qualities" and to focus researchers' efforts on practical means to find what explanations of nature they can from empirical observations and measurements. The challenges and pressures of the Cold War and space exploration in the 1950s and early 1960s motivated the development of new mathematical methods for dealing with complicated systems in which "noise" was a factor increasingly more crucial to deal with. The pioneering work is usually credited to T.R. Bashkow [Bashkow (1957)], R.E. Bellman [Bellman (1957)], and R. Kalman [Kalman (1960)]. Their work led to the development of what is today called "modern systems theory" based upon the mathematics of what is called "state variable theory." Much of this development was more or less an incremental outgrowth of older theories aided by the introduction of new mathematical formulations to overcome issues the older methods could not.

In the latter half of the 1960s another new paradigm for dealing more robustly with numerous practical engineering issues was introduced. This new paradigm, as it turns out, provides formal methods to put Slepian's Principle into widespread practice. It has come to be called the "set membership paradigm" or SMP. Its methodologies are called "set membership theory" or SMT. SMT first appeared in the works of Witsenhausen (1966; 1968) and Schweppe (1968). Its Cold War motivation is clear from the introduction section of Schweppe's paper:

> Processing *noisy* observations of some function of a dynamic system's state is often necessary to provide an estimate of the system's current state. The nature of the algorithm to be used depends on the assumed structure of the dynamic system, the observation errors, and the inputs driving the dynamic system. An algorithm is derived under the following three conditions:
>
> 1) The dynamic system is linear and completely specified, and the observations are linearly related to the state.
>
> 2) The input to the dynamic system is unknown except for a bound on either its magnitude or energy.
>
> 3) The observation errors are unknown except for a bound on their magnitude.
>
> Consideration of the preceding set of conditions is motivated by the desire to develop an estimator to track a target that is purposely performing evasive maneuvers in an attempt to prevent tracking. [Schweppe (1968)]

Schweppe's three conditions impose limitations on the kinds of problems his algorithm can be applied to. However, other researchers in the 1970s and 1980s were able to extend the scope of applications and expand it to deal with dynamic systems described by equations that are not linear. In the mid-1990s it was demonstrated that even the assumption that the system is completely specified (i.e., that its mathematical structure is known) could be relaxed [McCarthy & Wells (1997)]. In 2012 it was shown that the SMP was capable of being used to break cipher codes and decode encrypted messages, and that it was sometimes even possible for it to identify the author of an encrypted message [Carlson (2012)].

What all SMP methods have in common is that they produce *sets of possible solutions*, instead of only one "point" solution, which have the property that every solution is consistent with all the measurements that have been made plus everything known (or thought to be known) about the system *a priori*. Because every solution in the solution set is consistent with all known observed or measured data, *any* of these solutions can be used to describe the system. This, of course, is what Slepian's Principle states. The McCarthy & Wells (1997) paper even demonstrated that the method could identify when an *a priori* model of the system is inadequate to describe it. Although the SMP and SMT methods have been around for 50 years now, the community of scientists and engineers who know and practice it is still small and it is not widely taught in colleges yet because the community of practitioners is still small and its power to solve otherwise intractable problems in science and engineering has not become widely recognized. Yet, the ideas of the SMP and SMT methods are not difficult to grasp.

For the sake of specificity, suppose we are studying a system that we characterize as having: one observable (measured) property, y, (the "output"); one external measurable or observable property upon which the behavior of the system depends, α, (the "input"); and two parameters, x_1 and x_2, that determine a relationship between inputs α and outputs y according to some known or hypothesized relationship function f (the "model" of the system). The statement, "y is a function of x_1, x_2, and α" is expressed in the symbolic shorthand of mathematics as

$$y = f(x_1, x_2, \alpha).$$

The process by which an investigator comes up with this equation is called "modeling the system" or "the system identification problem" or simply "the modeling problem." When a mathematical form for the function f is the thing to be determined, this is called the *structure identification problem*; when quantitative values for the two parameters x_1 and x_2 are to be determined, this is called the *parameter estimation problem*. The modeling problem is usually broken down into five steps:

1) theoretical analysis of the problem;
2) experiment design;
3) structure identification;
4) parameter estimation; and
5) model validation.

If at step 5 the model is experimentally invalidated, the process is repeated until an acceptable model is found. The theoretical analysis provides what is called "*a priori* knowledge" of the system because this step usually precedes experiment and data collection. This "knowledge" is more properly called "*a priori* hypotheses" because if the underlying assumptions and basis in theory are faulty then so is the analysis. However, in the normal practices of science and engineering, scientists and engineers have a great deal of faith in their theories and so, to the extent the theory is correct, "*a priori* knowledge" can be applied to its outcomes. This sort of "knowledge" is what Kant called "contingent" knowledge, and all empirical theories consist of this kind of knowledge.

There is always uncertainty ("contingency") attending determinations of f, α, x_1, and x_2 because there is uncertainty attending every measurement and experimental result that led up to the ideas and theories that go into making determinations of models. What the set membership paradigm explicitly recognizes is that this uncertainty introduces *sets* of possible solutions for all four of these determinations. Some of these determinations can be held-to-be "more likely" than others but, even so, these others are not empirically *refuted* by the information the scientist or engineer has in hand at the time. In the future better instruments

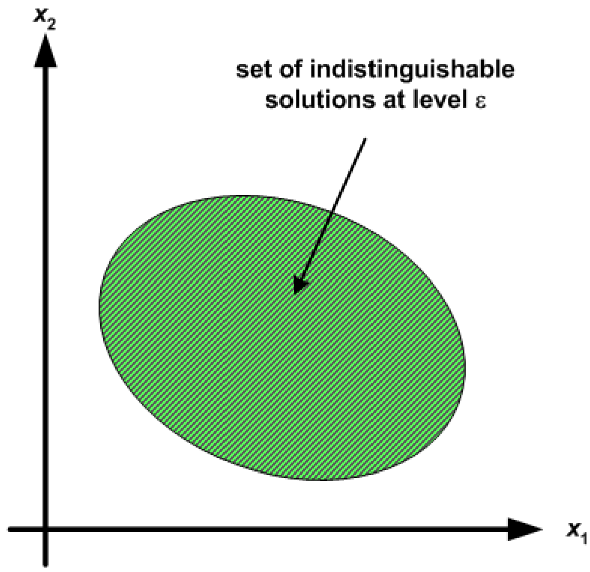

Figure 4: Empirically indistinguishable solutions of parameter estimates x_1 and x_2 at a level ε of measurement distinguishability.

and new discoveries might enable researchers to additionally rule out more of these "feasible solutions" and hold others of them "more likely" to be correct, but *at the time* the members of the feasible "solution set" all share the properties of 1) being consistent with all *a priori* knowledge of the system; and 2) being consistent with all empirically measured data that has been collected.

For example, suppose the problem is to estimate the parameters of the system. In general there will be multiple possible combinations of parameters that are feasible solutions to the estimation problem. Figure 4 illustrates this idea. Because all of them satisfy the two properties just named, they are empirically *indistinguishable* given the measurement capabilities of the instruments, the experimental data and observations made to date, and the existing state of the applicable theory.

In the illustration of figure 4, all the feasible solutions are shown grouped closely together in a roughly elliptical region of a "solution space" defined by x_1 and x_2.[3] Such a grouping is usually the case when our theories are confidently well established and have held up to scrutiny over an extended period of time. This "grouping behavior" of feasible solution set members is called the "Asymptotic Equipartition Property" of information, or AEP, by scientists who work in the field of information theory [Cover & Thomas (1991), chap. 3]. For practical reasons, scientists and engineers must usually *choose* one member of this set to use in carrying out their work. While *any* member of the set is an "acceptable answer," what people usually do is choose a particular solution from the geometric "center" of the distribution in the hope that this choice provides more "margin for error" in their application of the chosen solution. This can be called a *heuristic* for choosing because, in the past, choosing in this way has usually led to good results in their applications work.

Up to some limit, the more data we collect lets us eliminate ("rule out") more of the members of a solution space. More data is found with which previous members of the solution space no longer exhibit the key property of being *consistent with* our data. The same thing can happen when better measuring instruments are developed which provide less measurement uncertainty. In either case, the solution set is observed to "shrink" because more of its members are made distinguishable in the data collection process.

[3] It is called a "space" because we describe it in ways analogous to the way geometry was originally used to describe physical objects by the ancient Greeks. Because this "space" is mathematically abstract and the "points" in this "space" are solutions, it is called a "solution space." It is a metaphorical naming convention.

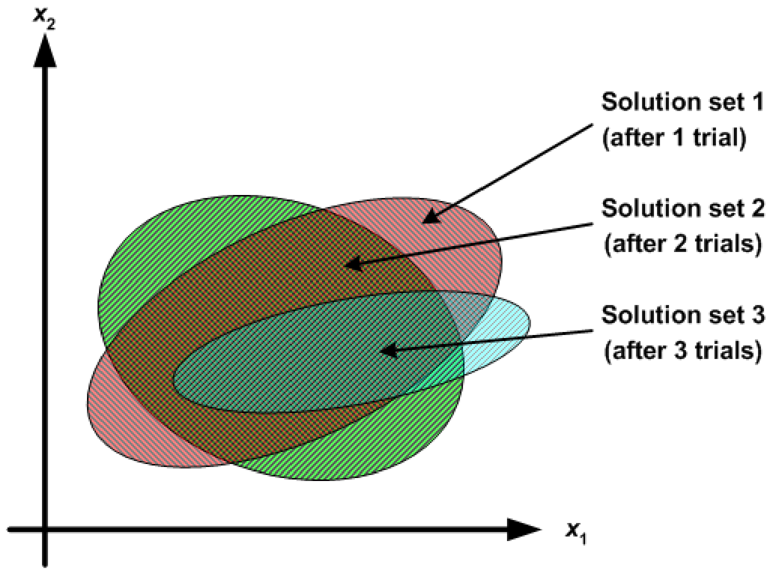

Figure 5: Reduction of the solution set through repeated data collection trials. After each successive trial, the feasible solution set consists of the mathematical intersect of the solution sets from each trial. Generally speaking, the "shape" of the solution sets gets more complicated with each successive trial.

Figure 5 illustrates this "shrinking" of the solution set. Typically the "shape" of the solution space gets to be more complicated with each successive collection of new data. A geometric description of the solution space becomes more complicated and difficult to express. This can oftentimes overwhelm the researcher's ability to provide an exact mathematical description of the solution space and so, for very practical reasons, he or she might resort to an *approximate* geometric expression that holds up at least near the "center" of the new solution set. The "method of optimal bounding ellipsoids" is one popular approach to this complexity problem. Another popular approach is called "outer-bounding by linear programming" [Belforte *et al*. (1990)]. Such methods can be likened to the method used by Archimedes to estimate the value of pi or to Newton's method of first and last ratios in his development of calculus.

As the number of data-taking trials is increased, one of the outcomes that can happen is that the solution set collapses to an empty or "null" set. (A null set is a set without any members in it; one example of a null set is "the set of all living dodos"; this set is empty because the dodo is extinct). There are two known reasons why this happens. The first is that the model structure represented by f is wrong – that is, the theoretical model f is not an adequately correct description of the system. For instance, the model f might be too simplified and fail to take into account dynamical aspects of the system essential for its correct description. The second reason is that the system being modeled might change over the course of time either in the parameters being used in it or because the structure of the system itself changes. For example, a buildup of ice on the wings of an aircraft changes the amount of lift produced by the wings and the model may fail to take this into account. Another example is provided by the phenomenon of metal fatigue, in which repeatedly flexing a metal part causes its molecular structure to change in a way that weakens the metal and can ultimately cause it to break.

One of the important contributions made by set membership theory is that its algorithms are capable of detecting when the solution set becomes empty. This property can be exploited by the researcher both in determining the model structure [McCarthy & Wells (1997)] and in tracking changes in its parameters that might be occurring over time [Rao & Huang (1993)]. At the same time, SMT algorithms are robust in situations where the model might have been made overly-complex – i.e., when it includes factors that are not relevant or do not pertain to a particular system. For example, consider two model structure functions

$$f_3 = x_1 \cdot \alpha + x_2 \cdot \alpha^2 + x_3 \cdot \alpha^3$$

and

$$f_2 = x_1 \cdot \alpha + x_2 \cdot \alpha^2 .$$

Suppose the system is actually governed by structure function f_2 but the person modeling the system uses function f_3 for his model. A set membership parameter estimator will return a numerical value for x_3 that is very nearly zero compared to the values it gives for x_1 and x_2. There are well known methods that exist for testing the significance in terms of the model's parameters, automatically "dropping" the x_3 term from f_3 and re-estimating the other two parameters in the absence of this term (f_2). This is called "model order reduction" and it can be made part of an SMT algorithm [McCarthy & Wells (1997)].

If the model structure function is adequate and the parameters remain fixed within a small range of variability, as more data trials are taken the solution set converges to a final set of consistent solutions. The number of solutions in the final solution set is an indicator of system uncertainty. It is the SMT analog to the "funny numbers $x \pm y$" discussed earlier.

One more remark is in order. In the discussion of SMT thus far, a tacit assumption lurks. This is the assumption that for a given system there is only *one* solution set. However, there are systems for which different and disjoint solution sets exist in different regions of the solution space. This would appear in figure 4 as a second solution set, disconnected from the one shown, in a different (x_1, x_2) region. For example, the physics describing a particle moving at a velocity very close to the speed of light are very different from those describing a particle moving at a velocity of less than 100 miles per hour. In the terminology being used here, the two systems have different mathematical structures. More commonly-occurring and "terrestrial" examples are often found in systems engineering where, for instance, "the" model is expressed by different structure functions dependent on a set of factors often called "the operating point" of the system. Electronic circuits employing devices called transistors make up one very common set of examples.

It is not especially difficult for the scientist or engineer to deal with such systems provided he knows he needs different model parameters for different regions of operation. But how does he come to know this? The "old fashioned" way of doing research in, say, engineering is to discover this by accident – sometimes very literally by accident as was the case when it was discovered that the aerodynamics of an aircraft change rapidly and sometimes violently when that aircraft "hits" the so-called "sound barrier." Discovery "by accident" of different solution sets for different operating regions is usually not a pleasing discovery for an engineer nor one especially helpful in advancing his career.

Fortunately, another technique exists that fits in very well with the SMP methodology. It was first reported in 1995 by Kennedy & Eberhart. They called this method "particle swarm optimization." PSO (as it is often called) is not widely recognized to be a set membership technique but the simple fact is that it employs solution sets and in no way departs from the set membership paradigm. I remember being very skeptical about PSO when I first heard about it. Fortunately for my personal edification, I had a young colleague who was an expert in this approach, and who demonstrated to me that he could use it to solve modeling problems that were beyond the capability of any method I knew at that time. Never did humble pie taste any better to me.

§4. Exploring the World of Facet B

A solution set can be regarded as another kind of "funny number" from the point of view that it is an answer to some specific question of empirical science or engineering. Like our "funny numbers of the class $x \pm y$," once we begin thinking about them as "numbers" there are many follow-on questions we might want to ask about them. Is it meaningful to add two solution sets? or to multiply two of them together? If so, how does one do this? What happens if we have two solution sets and they overlap (i.e., their mathematical intersect is not null)? Is there an algebra for solution sets? Can they be quantitatively compared and put into some kind of structured ordering? Why would someone even want to think about these and a host of other questions about "solution set numbers"?

In the earlier chapters, the point was emphasized that even so seemingly-simple mathematical entities as "numbers" are *invented objects*. They were invented to fulfill purposes of their inventors and allowed their inventors to solve problems that were for one or more reasons important to them. The objects of Facet B are **all** *noumena* beyond the horizon of possible human sensible experience, and their invention happened for practical reasons.

These and other historical observations regarding the invention of mathematics bring us back around to the question that began this treatise: How is it possible for mathematics to describe nature? We have come to the point in this treatise where we can now offer an epistemological answer to this question. *Mathematics describes nature* ***because we make it describe nature***. If a mathematical answer *fails* to describe nature, we go and find another mathematical answer that *does*. The very fact that human beings are able to conceive ideas of supersensible objects – to, in effect, *create* a whole world of *noumena* to invisibly accompany the sensible world of our experiences – is perhaps one of the greatest marvels of the human mind. Certainly it is this power humankind used to achieve dominance over every other species on earth, to tame forces of nature, to rise above the subsistence level of living characteristic of hunter-gatherers, and to make ourselves more than what we were the day we were born. Piaget wrote,

> Life is a continuous creation of increasingly complex forms and a progressive balancing of these forms with the environment. To say that intelligence is a particular instance of biological adaptation is thus to suppose that it is essentially an organization and that its function is to structure the universe just as the organism structures its immediate environment. . . . The organism adapts itself by materially constructing new forms to fit them into those of the universe, whereas intelligence extends this creation by constructing mentally structures which can be applied to those of the environment. In one sense and at the beginning of mental evolution, intellectual adaptation is thus more restricted than biological adaptation, but in extending the latter, the former goes infinitely beyond it. [Piaget (1952), pp. 3-4]

In ancient Greece the scientist and the mathematician were one and the same person and were called philosophers – lovers of wisdom. Archimedes of Syracuse is regarded today as one of the great mathematicians of antiquity; but the Romans regarded him as a formidable engineer whose weapons of war were responsible for holding them back from conquering Syracuse.

Gradually over time, a gap appeared and slowly widened between the scientist and the mathematician as some men focused their attention more exclusively on the empirical world of Facet A and others more exclusively on the *noumenal* world of Facet B. In the time of Newton, Descartes, and Huygens, this gap was still only barely perceptible. It widened in the 18th century, and in the 19th century it differentiated clear distinctions between men like Fourier, Faraday, Maxwell, and Planck in comparison to other men such as Dedekind, Cantor, Riemann, and Weierstrass. Today specialization is taken for granted as a matter of course and economic utility. However, the chief characteristic of specialization is deliberate self-limitation. Specialization does have certain advantages up to a point. Nobel laureate Eugene Wigner said,

> Physics does not endeavor to explain nature. In fact, the great success of physics is due to a restriction of its objectives: it endeavors to explain the regularity in the behavior of objects. This renunciation of the broader aim, and the specification of the domain for which an explanation can be sought, now appears to us an obvious necessity. In fact, the specification of the explainable may have been the greatest discovery in physics so far. [Wigner (1964)]

There is truth in what Wigner said. But there is also a problem inherent in it. If physicists no longer attempt to explain nature, then who does? There is a wise old saying that goes: "The specialist is a guy who knows more and more about less and less until eventually he knows everything about nothing. The generalist is a guy who knows less and less about more and more until eventually he knows nothing about everything." Does it not seem obvious that a proper balance between specialization and generalization lies somewhere in between these two extremes?

If this is so, then it reasonably follows that ability to make great contributions in science or engineering is partly contingent upon the ability of the scientist or engineer to also *be* a mathematician insofar as "mathematics" is regarded in terms of creation and design of mathematical objects and structures. It is sometimes said that "the job of the mathematician is to do proofs," but this is certainly not descriptive of all that mathematicians do in their day to day work. Poincaré wrote,

> What, in fact, is mathematical discovery? It does not consist in making new combinations with mathematical entities that are already known. That can be done by anyone, and the combinations that could be so found would be infinite in number, and the greater part of them would be absolutely devoid of interest. Discovery consists precisely in not constructing useless combinations, but in constructing those that are useful, which are an infinitely small minority. Discovery is discernment, selection. . . . Mathematical facts worthy of being studied are those which, by their analogy with other facts, are capable of conducting us to the knowledge of a mathematical law, in the same way that experimental facts conduct us to a knowledge of a physical law. They are those which reveal unsuspected relations between other facts, long since known, but wrongly regarded to be unrelated to each other. Among the combinations we choose, the most fruitful ones are those which are formed of elements borrowed from widely separated domains. [Poincaré (1914), pp. 50-51]

Clearly one is unlikely to be able to "form fruitful combinations" between objects "borrowed from widely separated domains" if one deliberately separates *himself* from those domains. A scientist or an engineer should recognize that some of the seeds of discovery lie in those "separate domains" that consist of knowledge of how mathematical structures and entities are first imagined and constructed.

Unfortunately, most instruction and training for scientists and engineers omits any mention that domains of mathematical invention and creativity exist. None can foresee any limitation to the mathematical world of Facet B, but it is easy to foresee that the greatest contributions to science, engineering, and technology of the future will be made hand in hand with mathematical creation and invention. Mathematicians are often thought of as discoverers, but to say this is to sell short their achievements. Mathematicians are also, and perhaps even primarily, *inventors*. This treatise might well be thought of as a journey of exploration into an unbounded world of imaginative invention in the land of Facet B. Let us undertake it as such.

References

Bashkow, T.R. (1957), "The *A* matrix: A new network description," *IRE Transactions on Circuit Theory*, vol. CT-4, pp. 117-120, Sept., 1957.

Belforte, G., B. Bona, and V. Cerone (1990), "Parameter estimation algorithms for a set-membership description of uncertainty," *Automatica*, vol. 26, no. 5, pp. 887-898.

Bellman, R.E. (1957), *Dynamic Programming*, Princeton, NJ: Princeton University Press.

Carlson, Albert (2012), *Set Theoretic Estimation Applied to the Information Content of Ciphers and Decryption*, Ph.D. dissertation, Moscow, ID: The University of Idaho.

Cover, Thomas M. & Joy A. Thomas (1991), *Elements of Information Theory*, NY: John Wiley & Sons.

Davis, Philip J. & Reuben Hersch (1981), *The Mathematical Experience*, Boston, MA: Houghton Mifflin Co.

Holton, Gerald (1952), *Concepts and Theories in Physical Science*, Reading, MA: Addison-Wesley Publishing Co.

Kalman, R. (1960), "A new approach to linear filtering and prediction," *Transactions of the ASME, J. Basic Engrg.*, ser. D, vol. 82, pp. 35-45.

Kant, Immanuel (1787), *Kritik der reinen Vernunft*, 2nd ed., in *Kant's gesammelte Schriften*, *Band III*, Berlin: Druck und Verlag von Georg Reimer, 1911.

Kant, Immanuel (1800), *Logik*, in *Kant's gesammelte Schriften*, *Band IX*, pp. 1-150, Berlin: Walter de Gruyter & Co., 1923.

Kennedy, J. & R. Eberhart (1995), "Particle swarm optimization," *Proceedings of the IEEE International Conference on Neural Networks*, IV, pp. 1942-1948.

Lakatos, Imre (1978), *The Methodology of Scientific Research Programmes*, Cambridge, UK: Cambridge University Press, 1980.

MacDonald, D.K.C. (1948), "The Brownian movement and spontaneous fluctuations of electricity," *Res. Appl. in Industry*, vol. 1, pp. 194-203, Feb., 1948, London, UK: Butterworth & Co.

Mandelbrot, Benoit (1977), *The Fractal Geometry of Nature*, NY: W.H. Freeman and Co.

Margenau, Henry (1977), *The Nature of Physical Reality*, Woodbridge, CN: Ox Bow Press.

Marías, Julian (1967), *History of Philosophy*, NY: Dover Publications.

McCarthy, S.G. & R.B. Wells (1997), "Model order reduction for optimal bounding ellipsoid channel models," *IEEE Transactions on Magnetics*, vol. 33, no. 4, pp. 2552-2568.

Piaget, Jean (1952), *The Origins of Intelligence in Children*, Madison, CT: International Universities Press, 1974.

Plato (date uncertain), *Republic*, *c*. 4th century BC.

Poincaré, Henri (1914), *Science and Method*, Bristol, England: Thoemmes Press, 1996.

Preparata, Franco & Raymond Yeh (1974), *Introduction to Discrete Structures*, Reading, MA: Addison-Wesley Publishing Co.

Rao, Ashok & Yih-Fang Huang (1993), "Tracking characteristics of an OBE parameter estimation algorithm," *IEEE Transactions on Signal Processing*, vol. 41, no. 3, pp. 1140-1148.

Schweppe, Fred (1968), "Recursive state estimation: Unknown but bounded errors and system inputs," *IEEE Transactions on Automatic Control*, vol. AC-15, no. 1, February 1968.

Slepian, David (1976), "On bandwidth," *Proceedings of the IEEE*, vol. 64, no. 3, Mar. 1976, pp. 292-300.

Smith, David E. (1923), *History of Mathematics*, vol. I, NY: Dover Publications, 1958.

U.S. Census Bureau (2011), *Statistical Abstract of the United States: 2012*, 131st ed., Washington, DC.

Wells, Richard B. (2006), *The Critical Philosophy and the Phenomenon of Mind*, available from the author's web site.

Whitehead, Alfred North & Bertrand Russell (1910), *Principia Mathematica*, in 3 volumes, available from Merchant Books, © 2009.

Wigner, Eugene (1964), Nobel Prize Acceptance Speech, Dec. 10, 1963. Reprinted in *Science*, vol. 145, no. 3636, 995 (1964).

Witsenhausen, H. (1966), "Minimax controls of uncertain systems," M.I.T. Electronic Systems Lab., Cambridge, MA, Rept. ISL-R-269, May, 1966.

Witsenhausen, H. (1968), "Sets of possible states of linear systems given perturbed observations," *IEEE Transactions on Automatic Control*, vol. AC-13, pp. 556-568.