International Journal of Theoretical Physics, Vol. 18, No. 8, 1979

Schrödinger's Cat

Elihu Lubkin

Department of Physics, The University of Wisconsin—Milwaukee, 53201

Received September 22, 1979

The issue is to seek *quantum interference effects in an arbitrary field*, in particular in psychology. For this I invent a digest of quantum mechanics over finite-*n*-dimensional Hilbert space. In order to match crude data, I use not only von Neumann's mixed states but also a parallel notion of unsharp tests. The mathematically styled text (and earlier work on multibin tests, designated MB) deals largely with these new tests. Quantum psychology itself is only given a foundation. It readily engenders objections; hence I develop its plausibility gradually, in interlocking essays. There is also the empirically definite proposal that (state, test, outcome)-indexed counts be gathered to record data, then fed to a "matrix format" (MF) search for quantum models. A previously proposed experiment in visual perception, which has since failed to find significant quantum correlations, is discussed. The suspicion that quantum mechanics is all around us goes beyond MF, and "Schrödinger's cat" symbolizes this broader perspective.

1. INTRODUCTION

1.1. Quantum Logic in Everyday Life was my first title—too dull. The central theme is application of the mode of comparison between theory and experiment developed in the quantum mechanics of the 1920's to investigations apparently remote from atomic physics. The quantum mechanical mode of comparison is defined in Section 8 on quantum logics and in Section 9 on the matrix format (MF). Some readers will prefer to start there. I define the logic as all the empirical probabilities indexed by the ingredients of the observer's experience, his states and tests, in contradistinction to common usage, in which the lattice of what I call sharp questions is singled out (but see von Neumann, 1962, p. 195).

The earlier sections sell quantum logic in everyday life. It is usually held that the crudely macroscopic has nothing to do with the epistemological complexity of quantum physics, indeed that a "classical world" (a muddy notion) free of ontological complication must exist for experiments to be possible. Particularly those trained to easily follow the arguments have also acquired this metaphysical prejudice. I have therefore felt it best to do metaphysics (Sections 1–7) before the theorems. This is faithful to my thought, which has grown from somewhat of a metaphysical miasma. Section 7 includes a notion distasteful to me, quantum logic with no internal observer.

In 1958 or 1959 I heard Niels Bohr lecture at Columbia about his suggestion that complementarity might illuminate psychology. I understood relativity of reality then. Yet I smirked, and did not recall Bohr's opinion until after the work of Section 6.3 and after the deformation of quantum economics into quantum psychology to be recounted. So I have myself shared physicists' incredulity, based mainly on the smallness of $\hbar$. This is not to insist that the essays are written only to sway doubters: the multifaceted approach of a series of essays is more suited to a voyage in complementarity than a more self-consciously linear style.

1.2. Schrödinger's Cat. The nontrivial structure of reality in quantum mechanics has a mascot. I first heard Schrödinger's fable (1935; Jauch, 1973) told without a moral in a classroom of T. D. Lee's. Once upon a time an unfortunate cat was put into a box together with a radioactive atom, so rigged that the decay would kill the cat. If the original and decayed atomic wave functions are a and b, the live cat's wave function c, the dead cat's d, then after a half-life, the development of an atomic state which if unlinked to the cat would have been $2^{-\frac{1}{2}}(a+b)$ instead results in the atom-cat state $2^{-\frac{1}{2}}(ac+bd)$. Two cat realities are contained in one wave function, and the impracticality of obtaining an interference effect from both terms may not prevent us in principle from considering such interference and therefore both realities together. (Perhaps d should not be quite dead if one is to speak of it as experiencing a reality! Einstein, 1949, will not kill a cat even in a thought experiment and puts a mark on a tape instead.) Plural reality fits a philosophy where the wave function or other formal structure of quantum mechanics is more primitive than reality, a scientific heresy discussed in Section 2. von Neumann unsuccessfully tried to squelch this plurality by warning his readers too pointedly against it (1955, pp. 413–421, especially observation by the "abstract ego" of a system which contains the observer's brain!). The emphasis on plurality is now commonly called the Everett (1957)–Wheeler (1957) or many-worlds viewpoint. My own similar understanding of quantum mechanics (1979) relates to some points in the present paper.

The dependence of our common experience of reality upon its genesis through Darwinian evolution within a structure not itself clearly real is a theme of Sections 3–6. Issues related to plural reality—whether interference is or is not in laboratory practice necessarily close to the world of few atoms, the relationship of relative reality to superselection rules and to complementarity, and a question concerning prior correlations in the theory of measurement—form Section 6.

1.3. The Program. MF requires the measurement of probabilities of different outcomes of each of several types of experiment, done by repeated trials of each type. Each experiment is considered as being done in two stages, the preparation of a state and the execution of a test. Eventually each state and each different outcome or "bin" of each test has an $n \times n$ Hermitian matrix associated to it, much as in conventional quantum mechanics. Limiting the broad scope defined in Section 8 to MF, Section 9.1, is discussed in Section 9.2. Picking n small so as to have few parameters: Section 9.3. What programs for fitting data by MF might be like: Section 14. We have written a program for $n=2,3,4$ (Lubkin and Lubkin, 1979b). Interest might be attached by some to the factor-analytic quality of such programs: Section 14.5.

1.4. From Economics to Psychology. When my wife studied economics, I spoke to her (Quantize economics!) about exercising the ideas of physics there. In 1970 I became troubled about two doubts. Quantum logic can be presented as a confrontation of a state by a test to yield a distribution of probabilities over outcomes (Section 8). Comparison between the observer as tester and the observer as part of a state tested is perhaps not an automatic feature of quantum logic so presented, but I wish to regard quantum logic as having such a feature. For observer A interacting with system B to be a new state system AB, presumably for testing by yet another observer, A and B must be sufficiently similar that their interaction in entity AB be describable.

The *first* 1970 doubt was this: In economics, A is an economist and B is a firm, country, or industry. For A and B to interact, the economist had to be a kind of firm, or the firm an economist—ridiculous. Yet by saying to myself for two years that the economist *is* a firm, I got used to the idea. To speed things up, imagine economist A theorizing about his own business with system B. A is now economic enough to be a firm.

Sad to say, this amusing symmetry between system and observer is not explicit in the main sequel, perhaps because I get tied down to adapting the limited state-test format to a pedestrian program of application. But

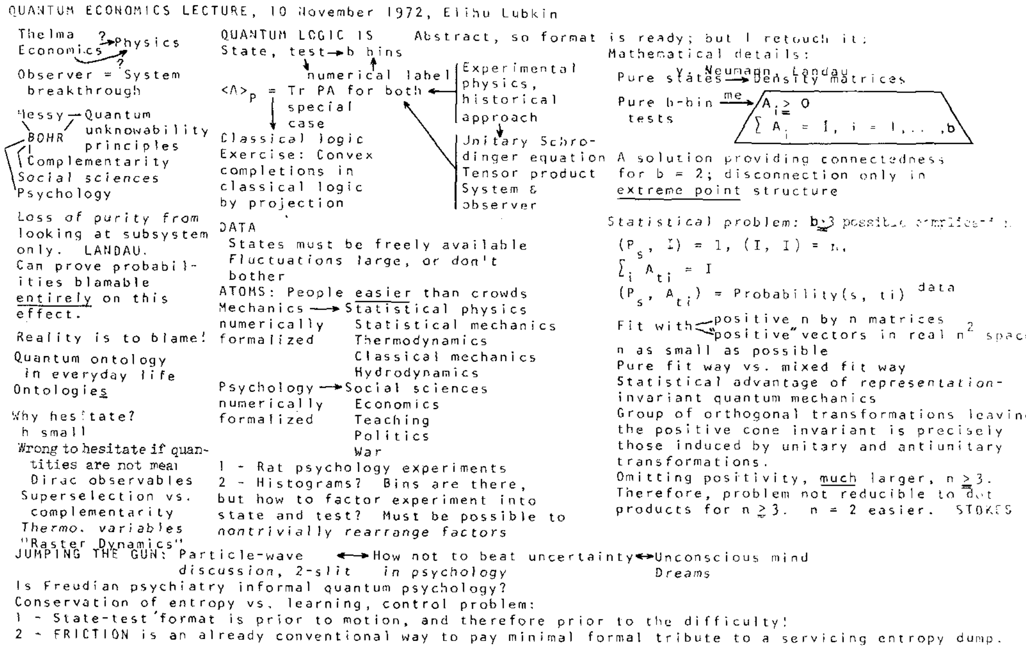

Fig. 1. Quantum economics lecture, 10 November 1972—a diagrammatic outline.

this program should lead to observers within states, if pursued. Then a quantum theory would make reality melt away. This brings us to my *second* 1970 doubt: It seems absurd that "interactions" of some crude economic model should melt physical reality. I reject this doubt because conventional physics is vigorous without absolute reality, hence absolute reality is discredited. Reality is open to rough treatment for the development of any phenomenology.

Many copies. Empirical quantum logic requires the duplication of many copies of its states, so much so that a state is defined in terms of the procedure for making one. Perhaps it is because one easily finds another atom that quantum logic has flowered in the atomic physicist's garden.

For an experimental economist to secure many copies of a "state," it is awkward to work with large firms. And interference may be masked by largeness. Instead of burying fluctuations under large numbers, we seek to probe fluctuations. So it is best to make the firm small: a person, a rat, a microbe: some economic atom. Later, "many-body" theory may relate fluctuations of small systems to features of large systems, as in physics.

The firm now is indeed like the economist—I hold this descent at the level of a person. But the economist has become a psychologist.

Quantum psychology.[1] It is easy to hope that experiments in quantum psychology may cement analogies between the vaguenesses of dreams and Freudian conceptions, and the complexity of realities in quantum theory. Introspective conjecture might most easily suggest experiments in human psychology, though rats or microbes would seem even more atomic. One might frame notions about thoughts of rats, to relate dream-analogy conjecture to animals: any crazy motivation is safe *for finding an empirical domain to study*, because MF would develop results mechanically from the probabilities measured by experiments, not from the motivating enthusiasms.

The intention is to abstract a simple quality as the entity described by quantum state and test matrices. It is like studying a spin or a few interacting spins, without worrying about virtual hyperons. But the phenomenology is to be psychological. The theoretical framework allows large interference effects. Should these appear, the concomitant relativity of reality in the new physics of thoughts would be a new departure from unified classical reality. This could act back upon general physics, where yet most thought remains saturated with classical reality borne by some

[1]Because there are people who indiscriminately connect quantum physics with the bizarre and who are independently enthusiastic about something they call quantum psychology, I feel obliged to make it known to the reader that I am not myself a devotee of either parapsychology, transcendental meditation, or religion.

space–time manifold (Lubkin, 1964). When it is not a question of framing formal structures, writers (e.g., Borges, 1967) have indeed sought to abandon classical reality.

1.5. Mixed Tests, Extraneous Tests. The state matrices P of MF are the familiar density matrices, but the bin acceptance matrices of Section 10 on tests (and MB), *acceptors* for short, were new to me. These generalize Dirac-observable tests in about the same way that density matrices generalize wave functions. The empirical program almost—the hedge is Section 14.3—demands this generalization, because it is not possible to decide a priori which experimental procedures correspond to Dirac observables. The generalization brings in convex combination of tests, hence mixed tests and extreme tests, the conventional observables being extreme tests of the sort I call sharp. I call tests not mixtures of sharp tests *extraneous*.

1.6. Notation. *Positivity* is strong. A positive matrix (Appendix A) is "positive definite," not a matrix of positive numbers. I is the unit matrix. $\lambda(M)$ is the eigenspace of matrix M belonging to eigenvalue λ, i.e., $(Mx=\lambda x)\Leftrightarrow[x\in\lambda(M)]$. Thus $0(M)$ is the kernel of M and $1(M)$ is the image of M. $\mathrm{diag}(a,b,\ldots)$ is a diagonal matrix.

Dirac bracket $\langle x|y\rangle$ is the Hermitian inner product of state vectors x and y, with $\langle\lambda x|y\rangle=\lambda^*\langle x|y\rangle$ and $\langle x|\lambda y\rangle=\lambda\langle x|y\rangle$ for λ a complex number; $\langle x|My\rangle$ is also written $\langle x|M|y\rangle$. The map $(A,B)\to\mathrm{Tr}\,A^\dagger B=A\cdot B$ from a pair of $n\times n$ matrices to complex numbers defines a positive definite dot product usually restricted to the n^2-real-dimensional vector space of Hermitian matrices, and written with a dot. Dot product of the traceless parts a,b of A,B considered as vectors in an (n^2-1)-dimensional space, is written (a,b).

When the $n\times n$ Hermitian matrices are considered a topological space, the norm topology in the n^2-dimensional vector space is meant: the set of matrices $M+X$ is an open sphere about M of radius ε, if X ranges over $X\cdot X<\varepsilon^2$, and these open spheres constitute a topological basis. Topologies of lists are direct-product topologies of those of the arguments; topologies of subsets are induced by restriction.

"Metaphysics" parallels "metamathematics." "Mysticism" denotes belief in absolute reality or, in Section 2, preoccupation with an underlying wave-function-like structure.

1.7. Related Work. References I learned about after writing the main text deserve early mention.

Greidanus (1971) is close in mood to the present work, though his stress on information theory is absent here, and the black-box quantum empiricism central here is absent there. This may exemplify a literature on the importance of subcellular entities to the understanding of consciousness. Our ignorance in spite of neurophysiological investigations is granted, and a science of purpose is demanded.

Schrödinger (1958: no cat) claims a religious orientation, yet his substantive text is happily free of religion. (In my lexicon, "religion" is pejorative.) Thus his discussion of "oneness of mind" mentions Upanishads, yet poses two real oneness problems: the agreement between experiences of different individuals, and the intuitive unity of mind of one individual in the face of a plurality of nearly autonomous control systems found in the structure of the brain—close to the baffling agreement of ontologies of my Section 2. To assert Oneness *and be satisfied* is religious; to *worry about how* a diversity congeals to onenesses is atheistic.

Another parallel to my thought is the notion of progress of physics away from space and time. But Schrödinger does get carried away:

> This means a liberation from the tyranny of old Chronos ··· [which] strongly suggests the indestructibility of Mind by Time.

And the slow progress of physics away from space–time is tagged "Science and Religion."

The arbitrary division of experience into object and subject, "objectivation," Schrödinger considers an attribute of the scientific approach alien to a true study of Mind, which must therefore be pursued unscientifically! The division of an experimental procedure into preparation of a state and execution of a test in this present work is a species of objectivation, hence taking Schrödinger literally, must lead nowhere as regards Mind. But since the division is (merely?) a matter of counting, assembling many experimental types from fewer states and tests, and with the same laboratory procedures allowed within the arbitrarily delimited "state" and "test" portions of an experiment, it is hopefully not that bad. Since the quantum ingredient entails ontological diversity (cat!) when tests happen within states, and so problems of oneness, it bears much similarity to questions about Mind.

I write in Section 4 that, though Darwinian mechanisms are not teleological, their summary is in a stupid way teleological. Schrödinger finds this reflection also useful (similarly!) but writes pages on "Feigned Lamarckism" to explain it.

Two references to Bohr (1934, 1950) given by Greidanus prompt remarks:

> ...the unavoidable influencing by introspection of all psychical experience, that is characterized by the feeling of volition, shows a striking similarity to the conditions responsible for the failure of causality in the analysis of atomic phenomena

(1934), combined with Bohr's understanding of "the failure of causality" as a positive program of trying quantum models, is like my suggesting an MF program for psychology. Yet it is hard to move psychologists by advertising a "failure"; indeed I find it hard even with a "program." The following (1934) is another example of Bohr's misleadingly defeatist tone, where he wishes instead to stimulate investigation:

> The strict application of those concepts which are adapted to our description of inanimate nature might stand in a relationship of exclusion to the consideration of the laws of the phenomena of life.

The negative tone projects Bohr's modesty enhanced by the absence of evidence.

Bohr (1950) bears an encumbrance:

> ...however far quantum effects transcend the scope of classical physical analysis the account of the experimental arrangement and the record of observations must always be expressed in common language supplemented with the terminology of classical physics.

The supplement of classical terminology, especially canonical formalism, is a dark mystery which almost vitiates the clarity the statement had had had it ended with the words "common language." In atomic physics, the common language *was* classical canonical formalism plus electromagnetism, but Bohr's implied advice that canonical formalism must be introduced along with some model theory before experiments can be analyzed in regard to complementarity, or "quantum-logically," is an impediment that I have disregarded, except to add it now to the "luxuries" of Section 10.5. Bohr himself subtly doubts the need for "terminology of classical physics" by calling it a "supplement" to "common language" proper!

Bohr remarks (1950) that commutativity of the difference of coordinates $q_1 - q_2$ of two mechanical systems with the sum $p_1 + p_2$ of their momenta does not circumvent complementarity. This is close to Section 6.3. A searchlight is nearly cast upon the relativism and ontological complexity introduced through the essential neglect of the laboratory background for an observation, by Bohr's seemingly pedantic remark.

The breadth of Bohr's notion of complementarity even reaches (1950) vaguely towards metamathematics in his contrast between the practical use of a word and attempts at its definition; compare my Section 16.3.

My final reports of late browsing refer to Colodny (1972).

Arthur Fine dislikes the term "quantum logic" because ordinary logic still governs our writing; Finkelstein disdains such inhibition to stress the primacy of experiment. The subtlety in the dual meaning of "logic" is hinted at in Finkelstein's Note 1, p. 65. My own clarity is Section 8; subtlety, Section 5.

Finkelstein devises classical models for quantum logic. Does this detract from my Section 3 on switching mechanisms?

C. A. Hooker hopes that the epistemological development occasioned by quantum mechanics may yet again affect particle physics. The usual synthesis of the world from particles in space–time is physics "from the bottom up"; physics "from the top down" takes seriously the epistemological priority of experiment however gross the apparatus seems as an assemblage of Democritic particles. And...

The world seen as a connected whole "from the top down" might also furnish new insight into living processes (an emphasis Bohr also stressed).

Feinberg's article on particle democracy could be stretched to a claim that the "bottom" and "top" approaches coalesce, an observer being a large particle. If this is stretching a point, that may be what is needed to generalize space–time.

Colodny's *Introduction* quotes Hermann Weyl (1949, not 1926)[2]:

> It must be admitted that the meaning of quantum physics, in spite of all its achievements, is not yet clarified as thoroughly as, for instance, the ideas underlying relativity theory. The relation of reality and observation is the central problem. We seem to need a deeper epistemological analysis of what constitutes an experiment, a measurement, and what sort of language is used to communicate its result. Is it that of classical physics, as Niels Bohr seems to think, or is it the "natural language," in which everyone in the conduct of his daily life encounters the world, his fellow men, and himself? The analogy with Hilbert's mathematics, where the practical manipulation of concrete symbols rather than the data of some "pure consciousness" serves as the essential extra-logical basis, seems to suggest the latter. Does this mean that the development of modern mathematics and physics points in the same direction as the movement we observe in current philosophy, away from an idealistic toward an "existential" viewpoint?

This coincides with my criticism of Bohr. My Section 8 presupposes "natural language." Comparison with metamathematics dimly ties in with Section 16. "Existential" is prefatory to...

2. A MIX OF EXISTENTIAL AND MYSTIC VIEWPOINTS

An ontology develops from experience. This "existentialism" *seems* to awkwardly take macroscopic structures, especially people, as more basic than atoms. Epistemology vs. physics!

[2]The quote is from Appendix C, *Quantum Physics and Causality*. Analogy with *The Structure of Mathematics*, Appendix A, is hinted at by Weyl more than once. Appendix E, *Physics and Biology*, is close in spirit to the present article.

The common "mystic" resolution is that, though experience is the *source* of knowledge, the knowledge gained reveals a unified underlying absolute reality, most sharply drawn in the Newtonian view of particles moving in one space–time, somewhat blemished by a Copenhagen viewpoint. This *seems* to place people and other complicated junk in a properly subsidiary way.

My point here is that it is the other way around: The existential attitude takes *greater* cognizance of the awkwardness that we are our own "nonelemental" channels of information than does the mystic attitude, not less. The universal reality of the mystic is existentially suspicious partly because it also is built *from the experience of complicated junk*. From this existential viewpoint, experience in humans induces ontologies, with parallels in other forms of junk. The induction of an ontology is worthy of study; agreement among ontologies in forming a rich and stable practical "real world" is baffling, and is a topic. Mystic refusal to entertain such questions is unreasonable except in deference to a lack of ideas.

Out of this baffling agreement of ontologies has nevertheless somehow come a mathematics and a science. It is through this science that we judge ourselves to be complicated junk in the first place, so that it would be inconsistent to refuse the structures of physical theory a prominent place in this balance of attitudes. Perhaps physics will refurbish the mystic.

The perhaps mystic form I prefer for the placement of physical theory is to imagine an underlying physical structure. Experience of junk leads to distorted glimpses of this underlying structure, which yet reveal much about it. A particular ontology is, however, so distorted that in spite of the baffling agreement between ontologies, any one individual's or committee's "real world" is not even a model of the underlying structure. So the underlying structure is not "real." Realities over this structure are nevertheless of merely secondary scientific interest, the underlying subreal structure itself being of primary interest. Reality is not neglected, but is placed in a proper subreal context.

This personal "mix of existential and mystic viewpoints" was implicit in my taking the *wave function* as fundamental in learning quantum mechanics, was thoroughly ingrained by reading von Neumann, and is an irreversibly stubborn "Copernican" process of thought. The multiple realities in quantum mechanics degrade any one reality to an incomplete status congenial to the Berkeley–Hume–Borges analysis (Borges, 1967). Just as the earth is no longer a significant center, so now the experience and reality of one perceiver is not central in fixing the fundamental, underlying structure.

It may be well to rephrase here: Science teaches us that perceivers are complex, and leads us to frame our attitudes around simpler atomic constituents interacting over ordinary space–time. But the same distrust of complexity of perceivers can grow into distrust of the reality seen by the perceiver, perhaps not so much in regard to the atoms as in regard to the space–time in which they are to be housed. The very distrust of complexity which tends to induce a hard Newtonian mystic real 3-space for the sake of the particles, becomes existentialism when carried further. This existentialism stems not at all from humanistic promotion of human perceptions over scientific simplicity, but rather from the reverse, a distrust of all perceptions except for a scientifically simple extract. The psychology of perception becomes important for the foundations of physics, so that we may analyze our intuitive notions of reality in order to see what portions may be cast out of the underlying theory.

3. CONSCIOUSNESS IS PROBABLY NOT A SWITCHING MECHANISM: WE MAY BE ALMOST ONE-CELLED ANIMALS AFTER ALL

In order for an interaction to convey a property of a system observed to the memory of the observer, it is almost necessary for the interaction to have a *feeble* effect on the system observed. Otherwise the value of the information conveyed is likely to be vitiated by modification of the system observed in the process of measurement.

Control, on the other hand, implies significant modification. The exigencies of optimal control will likely dominate over keeping interactions weak to help build a classical ontology.

Darwinian evolution of optimal control is therefore probably *incompatible with a classical ontology.*

A switching system produces definite output for definite input, except perhaps for undesirable noise; one does not expect the intrinsic noise of quantum mechanical uncertainty. Therefore, it is unlikely, whatever the mechanisms of consciousness or purpose—the main control mechanisms—are, that they are a switching system. Moravec (1979) presents a contrary opinion.

Yet where the interactions are sufficiently gentle that they convey stable information, it is useful in a Darwinian sense to be able to store and process this information in a switching system. Such a peripheral switching system is likely, owing to its elaboration, to be the most obvious part of the whole mechanism, without being the whole thing.

If one relates this to structure of nerves, it is plausible to relate switching to interneuronal transmission, and the essential control mechanism to mysteries within a cell. One-celled organisms do have control mechanisms, with Paramecium (e.g. Satir, 1961) even having switchlike mechanism within its one cell. Once a control mechanism appears, the elaborations are likely to be in the peripheral switching system. Dictators persist.

Perhaps also, notwithstanding Section 6, there is after all some sense in associating quantum logic with the very small. Ingress into quantum coherence effects by means of gross psychological experimentation could amount to using a living organism to magnify phenomena of the very small, much as atomic physics experiments do, but with equipment developed by Darwinian evolution rather than by art.

4. QUANTUM ONTOLOGY AND FREE WILL

Remark: Approximations. Those things understood as approximations are those things regarded as truly understood. The fundamental level of questions has bypassed such things; they are no longer mysteries. In the sense that an approximation is "wrong," those things that are truly understood are those things known to be wrong! The attempt to understand something else is the attempt to show it thus to be wrong. Quantum logic shows classical ontology to be wrong, yet sometimes approximately right. Through this very attrition of its fundamental quality, ontology has become a thing better understood, malleable. The inexactness of its classical image becomes subject to investigation, which was not true when reality was perfectly sharp.

Consciousness is one of the "weapons for survival" induced by the mechanisms of Darwinian evolution in living forms. Control requires frictive dumping of entropy. Maximizing strength of interaction where control requires strength turns on quantum ontology. The subjective aspect of branching is not branching itself, but random behavior (Lubkin, 1979), nonzero lower "uncertainty" bounds on probabilities. This approaches "free will," yet does not arrive: each branch of reality, after an event involving an irreducible quantum element of "choice," is "occupied." Straightforward free will would have the will choose which branch will become real, or would at least improve over chance. So far the "free will" is no better than noise.

It is the part played by control, or Darwinian evolution, which does better than noise. Although the mechanisms of Darwin are not teleological, their summary is in a stupid way teleological. Consciousness and ontology are tautologically inseparable. The evolution of consciousness is therefore also evolution of ontology. The effectively teleological quality of evolution imputes a teleological quality to ontology. This is saying that that world view is evolved which is "best for survival." But since reality is an aspect of the world view, not a preset pattern within which the world view evolves, the world itself is evolved to "best" fit the organism.

Now let us shift attention to the day-to-day conscious thought of one organism. The formal aspect of control, the development of purpose, goals, learning, as a frictive process leading to a decrease of entropy in the sector of concern, is like Darwinian evolution. This is control along with a calculus of purpose, as distinguished from control in a conventionally engineered device slaved to a fixed "purpose" extraneously injected. Therefore, the fabrication and evolution of ontology in 2-billion-year evolution could seep into a theory of consciousness, if the mind gets described by matrices of quantum logic, if the elaboration reaches to the description of frictive control, and if it somehow grasps that aspect of the state of mind that expresses the world view. One cannot go further in free will than such bending of the reality of the world to fit the teleology of will.

The guess that somehow quantum ontology could transcend the inability to choose a branch of reality, and thereby achieve a nontrivial free will, was ineffectively told to me by Henry Stapp ~1961. See also Lucretius (Latham, 1951, pp. 66–68).

A discussion of free will by Martin Gardner (1973) prompts further words.

There are two aspects of "free will" so different that it perhaps seems arbitrary to lump them. One is a mere lack of predictability. Whether unpredictable mechanisms exist within the mind is more interesting; in order to try to define them one can exclude coin-tossing from free-will experiments.

The really interesting aspect of free will is effectiveness of control. Even if of the alternatives it is clear which one is best selected, there is the question of whether reality will be bent that way. I have suggested that control and noise are inseparable, strength of control introducing quantum logic, hence, uncertainty principles. "Free will" almost says this anciently: "free" signifying arbitrariness hence noise, "will" signifying control.

In Gardner (1973), a superbeing is posited who can predict which of two choices a player will take. Oddly enough, this superbeing is *not* described as a Poker Master. If the player's probabilities for dropping into the two bins 1, 2 are calculable from amplitudes a, b in a quantum psychology to be $|a|^2$, $|b|^2$, with $|a|^2 > |b|^2$, then the most astute Poker Master could predict the outcome no better than a fraction $|a|^2$ of the time, and he would do this well only by always betting on bin 1. This is because both branches of the wave function coexist coherently for possible tests other than the 2-bin test of the experiment.

Some place for quantum effects in psychology may hopefully be discovered in experimental domains built up out of fast and banal responses (e.g., Section 15.2). But the quantum effects should be volatile in matters *described as* conscious decisions. A "decision" based upon indoctrination is predictable. Only among the nebulous *original* decisions should quantum effects be manifest. It is hard to be repetitiously original; quantum psychology experimenters face a design problem in seeking intelligence while avoiding learning.

5. HOW CAN ONE PURPORT TO THINK IN CONFLICT WITH LOGIC?

All the arguments here are meant to be understood in a conventionally logical way. "Quantum logic" is for *relating theory to experiment*. Insofar as this may be kept separate from theory, there is no inconsistency. The reduction of experimental counts in bins of tests by means of MF, Section 9, is indeed as ordinary a calculation as anything else done on paper or in computers; while quantum logic shapes the theoretical half, it is not there in the metalinguistic rules.

But insofar as our language is also part of experience, there should be no natural wall between the two. If there is no natural wall between a nonclassical logic of events and a classical logic of our expressed thoughts about these events, then there must be an artificial wall, and there is: Greek axiomatization. This image of an axiomatic system as a wall between experience and further language is commonplace. More on the wall in Section 16.3.

The emphasis on a new "logic" is serious, while paradoxically not impugning traditional reasoning. That thinking be *written* out in new rules of "quantum logic" is not advocated. But when thought is the *object* of study, quantum correlations may be revealed.

The analogy between the wall of axiomatization and the cut between system and measurement device (von Neumann, 1955) will probably have occurred to the reader.

6. WHAT IS THE CONNECTION BETWEEN SMALLNESS AND ONTOLOGICAL INTERFERENCE?

6.1. Complexity vs. Interference. The following standard theorem of ordinary quantum mechanics gives sufficient conditions for interference to be *un*detectable; it describes how messy coherence looks like incoherence. Objections to the possibility of quantum mechanical interferences in psychology based upon it make it and its proof relevant.

Theorem 1: Interference Theorem. If states x and y contain orthogonal attributes that are not involved in an observation of $\alpha x+\beta y$, then that observation cannot distinguish between states $|\alpha x+\beta y\rangle\langle\alpha x+\beta y|$ and $|\alpha|^2|x\rangle\langle x|+|\beta|^2|y\rangle\langle y|$.

The meaning of "orthogonal attributes" is that $\langle x|A|y\rangle=0$ for the observables A (or acceptors) associated with the procedure of observation or test (Lubkin, 1979). This is most often known when there is a tensor-product structure with $x=x_1\otimes x_2$, $y=y_1\otimes y_2$, and where $\langle x|y\rangle=(x_1,y_1)(x_2,y_2)=0$ because both (x_2,y_2) is known 0 and $A=A_1\otimes I$.

Proof. That the two states are indistinguishable follows from $\mathrm{Tr}|\alpha x+\beta y\rangle\langle\alpha x+\beta y|A=\mathrm{Tr}(|\alpha|^2|x\rangle\langle x|+|\beta|^2|y\rangle\langle y|)A=|\alpha|^2\langle x|A|x\rangle+|\beta|^2\langle y|A|y\rangle$, owing to the given vanishing of the "interference" terms $\langle x|A|y\rangle$ and $\langle y|A|x\rangle$. ■

The interference theorem may be paraphrased: Only states *simply* different may visibly interfere. If an observation is to sense the relative phase α/β, then x and y must be simple, or if not, their complicated attributes must coincide or at least be at an angle in Hilbert space significantly different from a right angle.

If x and y are states of a psychological or sociological system, they are very complicated from the point of view of atomic physics. States worth labeling *distinctly* owing to properties discernible through psychological tests may be expected to have complicated atomic differences, and so be incapable of demonstrating interference. This would leave quantum psychology dead.

The most direct possibility yet open for interference is a psychological difference which *is* simple even on the atomic level. If the dual evolution of control and of ontology under the Darwinian stress of optimization indeed requires ontological interference at the center of control, and if interference phenomena cannot be atomically complex, then the center of control must be atomically simple. If so, the volitional acts of living forms, evidently an amplification of "thoughts" involving less matter and motion than the consequent "acts," have been amplified all the way up from the atomic level. This would be analogous to the amplifications that constitute the laboratory procedures of atomic physics, in building differences simple on the atomic level into "macroscopic" differences.

But in a formally *new* description of a psychological system by a state space, the atomic viewpoint need not be recognized; the states may be *formally* simple without regard to the above. The notion that *broad* interference phenomena are therefore possible arises. This may be unduly simple-minded, like expecting two-slit diffraction to work while one slit is being monitored just because one has not imagined how to introduce the monitor into the theory. However, the criticism that complexity precludes interference follows from the tensor-product construction of the complex from simple parts. This criticism becomes unconvincing when the atomically complex system is newly invested with a state-space structure not referent to atomic structure at all.

Whether interference effects in psychology are subtle in being truly simple atomically yet being demonstrable by amplification outputting psychologically, or whether interference effects are gross, indeed whether they exist at all, is a question for experiment. If subtle, then experiments will have to be fast and careful. The psychological states x and y must not be distinguishable by something as "classical" as even one neuron's potential; one must test $\alpha x + \beta y$ by amplification before the intervention of any competing amplification which threatens to convert the originally atomic "states of mind" x and y into objects distinguishable by means of electric potentials across cell membranes.

"Raster dynamics" poses a hope for gross ontological peculiarities.

6.2. Raster Dynamics. I return to our mascot in order to show that the argument for ontological branching is so general as to possibly apply in physical contexts not usually associated with quantum mechanics.

The evolution of system's state x_i together with observer's state y_o, ready to observe, to a combined state wherein the observer has learned about the system's state, is denoted $x_i y_o \to x_i' y_i$; see however Section 6.3, *Caveat*. "x_i'" instead of x_i provides for some perhaps slight modification or recoil of the x system's state in response to the interaction. Then the linearity of motion and the possibility of preparation of the coherently superposed x state $\Sigma_i a_i x_i$ imposes ontological branching upon the observer or joint xy system: $\Sigma_i a_i x_i y_o \to \Sigma_i a_i x_i' y_i$. No one outcome i is chosen; all appear together in a coherent sum. If x_1, x_1' are undecayed radioactive nuclei and x_2, x_2' are decayed, y_1 a live cat, y_2 a dead cat, $i = 1,2$, we have the ontologically branched cat of Schrödinger. Meow.

The argument uses linearity of motion. In this regard it appears applicable not only to quantum mechanics but also to any theory where, possibly in some approximation, there are linear equations of motion, in particular to the theory of small vibrations in classical mechanics.

The argument also uses tensor product for putting the x system together with the y system to form a compound system, and treats the observer as a y system: The observer lies within the domain of validity of the linearity of motion and of the rule that combination of systems be effected by multiplication of their representative linear-space elements or "wave functions." In this regard, the argument still fits quantum mechanics considered as a theory that ought in principle to encompass states of the observer but does not fit ordinary classical theory of small vibrations, the observer being neither usually regarded as classical, nor small, nor a vibration.

It is amusing to seek nevertheless to impose the conclusion of ontological branching upon classical mechanics, at least within its small-vibration sector, by seeking a special context wherein the unusual extension of linearity of the xy complex and the multiplication of systems are valid.

For this purpose, I have for a long time had a vague picture of a complex "vat of jelly," constituted of ordinary matter but being described in a classical manner, at the level of phenomenological engineering, in terms of linearly superposable waves. The waves are gross structures like water waves, the wave functions are ordinarily "real" to the engineer using his eyes to see them, not subtle tools of atomic physics. What is vague is that the jelly is to describe at least an xy complex, not merely the x system, and that the normal modes naturally be indexed by pairs (i, j) of integers so that the (i, j)th wave $(xy)_{ij}$ is roughly a product $x_i y_j$ of simpler waves x_i, y_j in "factor" systems. Although "vat of jelly" fails to describe the factoring structure, that factoring structure is a specialization. What remains to be done in rendering the classical mechanism vivid is details of how to engineer a specialization.

About "factorization": The xy system should have its state space linearly spanned by the $x_i y_j$, there should be approximate motions $x_i(0)y_j(0) \rightarrow x_i(t)y_j(t)$, t denoting time, where xy joint motion factorizes into separate x motion $x_i(0) \rightarrow x_i(t)$ and y motion $y_j(0) \rightarrow y_j(t)$, but the exact motion should not factorize, this lack of factorizability of the joint motion constituting the interaction.

We also wish to regard the y system as enough of an observer to have a point of view or at least an evolving memory storage. Otherwise, the combination of different y_i in state $\Sigma_i a_i x_i' y_i$ would not bear the character of *ontological* splitting. The reader will not be surprised that the injection by construction of a mechanism for consciousness in the y system has not been accomplished or even attempted by the writer.

But imagine all these things done. Then, even while an ordinary engineer watches the waves move in his vat of jelly, within that vat there is an observing y entity which undergoes ontological branching. For us to examine a reality for this y entity, a "y viewpoint," we must select one branch. In the y viewpoint, the results of experiments on the x system can be anticipated only statistically, the pure or mixed state of the x system is described by a wave function or density matrix associated with a method of preparation in such a way that the state may be known owing to prior statistical study of states similarly prepared, but not directly and not even by repeated trials of only one single method of preparation: There is a whole quantum mechanical family of relative realities associated with the y system factor of the vat of jelly where not only are the realities mutually unreal, but they are of a different nature from the overall reality seen by the external engineer.

"*Raster dynamics*" is a program to replace the word "jelly" by something closer to the desired structure of tensor factorization, yet without attempting y memory: Imagine a rectangular array or "raster" of masses at each intersection of I horizontal lines or "rows" and J vertical lines or "columns." The IJ masses interact through some complex of springs so arranged that it is possible to "bow" the whole raster with *row frequencies* which excite all the rows so similarly that the column linkages are not activated, the rows moving in unison, with a motion therefore describable in terms of the motion of a simpler system, one row. There are to be a distinct set of *column frequencies* for bowing the columns in unison without activating the row linkages; these reduce to one column.

I suspend further invention of the raster; the nature in which the factorizability is "classically modeled" should be already clear. Of course factorizability should be inexact to depict interaction.

The usual picture of a "jelly" would involve a single 3-space vector displacement of a 3-space argument, at any one time. Another factor could be introduced by arranging for a classically intricate "unit cell," leading to Rayleigh-like "optic modes." This "jelly" is now like the "raster."

It is well known that quantum mechanics can be considered as contained within the small-vibration theory of classical mechanics. My hope that nonlinear corrections within a truly classical theory of moderate vibrations could be followed within an internal y system's quantum ontology, in order to present linearity in quantum mechanics as a smallness approximation, was Lubkin (1968). The difficulty of marrying nonlinearity with "tensor product" has proved too confusing for me. But it seemed that, for a slightly nonlinear Schrödinger equation, the tree structure of an internal observer's ontology remains, in spite of the possibility of sharper experiments which support questions that refine the lattice of sharp questions from subspaces of Hilbert space to subsets. The duller experiments still lead to a quantum ontology if the observer himself, as is intended in "raster dynamics," is built out of small vibrations. An independent judgement to this effect appears in Mielnick (1980).

The engineering of a raster *here* is intended merely to reveal the internal ontological complexity of a linear classical system, not expressly for nonlinear generalization.

So far I have sketched a raster in our own standard world of atomic physics, with our own ontology within the branching structure of ontologies proper to the ordinary quantum mechanics of polyatomic systems, and with an ontologically disparate branching structure to describe the realities of the y system within the raster. The y realities "within" the raster are not for us possible realities: Though we may build other worlds, we may not live in them, being ourselves stuck in the world of atomic physics.

In the sequel, it is argued that this clarity of delimitation of our own "ontological range" is *not* with certainty that sharply drawn, though we cannot expect to directly identify with the y system of an engineered raster. As a further prelude to this goal, it is instructive to consider a second example of ontological complexity, associated with a debate about superselection rules.[3] I make the point telegraphically in Lubkin (1971), and in its title, *Accepting Superselection Entails Rejecting Complementarity*. For clarity I redo this under a new slogan:

6.3. Quantum Mechanics with Complementarity is Dirty. The dirt is the unmonitored transfer of one quantity q, measurable but not being measured, to a system being prepared to have a definite value for a complementary property p. This physically dirty aspect underlying the deceptively precise theory of state vectors in a Hilbert space was clear to Bohr. One learns about complementarity, superselection, and also about ontological relativity—Bohr's complementarity encompassed a good deal of ontological relativity.

The system of interest is system 1. It is studied by a laboratory, system 2. The 1,2 complex is, in turn, studied by another laboratory, system 3. Whether some "4" looks at 3 or not will not specially concern us.

Let q be additive and conserved: $q_{1,2}=q_1\otimes 1+1\otimes q_2$, and if 1,2 is isolated, the value of $q_{1,2}$ is a constant of the 1,2 motion.

Let the 1,2 laboratory initially be in an eigenstate of $p_{1,2}$, the operator for an *additive and conserved complementary property p* appropriate to the 1,2 system: complementarity here refers to noncommutativity of p and q operators, not to more particular commutation relations. Indeed, let this be because the 1 system has p value p_1' and the 2 laboratory, p value p_2', so that $p_{1,2}'=p_1'+p_2'$ is the value of $p_{1,2}$ as seen by 3.

Next, within the 1,2 system, 2 measures the q property of 1 as defined by 2 while 1,2 remains isolated from 3. The 1,2 state is now of special q_1 eigenstate correlated form $\Sigma a_i x_i y_i$, where $q_1 x_i = q_{1i}' x_i$ and y_i is the 2-laboratory with a record that the 1 system has q_1 value q_{1i}'. The overall 3-seen p

[3]Wick, Wightman, and Wigner introduced the superselection rules (1952). They are called into question obliquely in Lubkin (1960). Clear statements are Aharonov and Susskind (1967a, b) and Lubkin (1970). Also see Epstein (1960), Rolnick (1967), Hegerfeldt, Kraus, and Wigner (1968), Wick, Wightman, and Wigner (1970), Mirman (1969, 1970, 1979), Lubkin (1977).

value is still $p_1'+p_2'$. However, the state x_i relative to the 2-observer y_i, in being a q_1 eigenstate for 2, is not described as a p_1 eigenstate by 2. The mix of 2's p_1 eigenvalues involved in 2's state x_i, on which q_1 assumes value q_{1i}', is in contrast with 3's sharp value p_1' prior to the 2-measurement of 1. If 2 and 3 agree as to p values of 1, then the 2 measurement has involved an unmonitored transfer of conserved quantity p between the laboratory 2 and the system 1—unmonitored else the amount of change of the p_1 value would be known and the changed p_1 value would be sharp, but it is not sharp. If p is momentum and 2, 3 are both massive, then 2's recoil *velocity* is negligible, and 2, 3 *will agree* as to p values of 1.

Hence, in order for the 2-laboratory to recognize both p and q eigenstates of the 1 system as legitimate in the state space of 1, the 2-theoreticians must tolerate uncontrollable transfers of quantities between their laboratory and the system 1, at least in the process of forcing the system-1 state from one definite eigenstate (of p) to another (of q). Although system 1 has associated through all this a Hilbert space with clearcut states before and after, with the 2-system formally in the background for the 2-theoretician, this is only in virtue of a theory willing to tolerate 1–2 contacts that allow unmonitored transfers of quantities, contacts like those of an open thermodynamic system. The theory requires dirt under the rug!

It is I think because this can feel sloppy that superselection rules are favored for some quantities like electric charge. Such superobservables commute with all physically legitimate density matrices. There is no complementary measurable quantity; a physically achievable *pure* state must be an eigenstate of charge.

An obvious violation is a grounded conductor. The charges here lack the empirical individuality of particle tracks in a photograph, that is the requisite sloppiness. It is not clear that the conductor has a definite number of electrons, even if the state is the well-defined ground state at absolute zero. The ground is part of the laboratory, system 2, and the charge on it does not count; only the charge on 1, the conductor, counts, the quantity being uncontrollably transferred being electric charge. What "the charge" is depends on where the conductor ends. If it is delimited without regard to charge and attached to a "ground" not part of the system of interest, then a state of conductor and ground with definite *net* charge will be a *coherent* mix of terms with different values of *conductor's* charge; this is made plain by the familiar coherent sharing of an electron in a bond between two atoms.

This coherence will become lost to us if we say, the ground itself is so uninteresting that we replace the density matrix of conductor and ground together by the reduced density matrix obtained by trace on the ground's labels: *Landau tracing*.[4] Such would be appropriate to predict expectation values of observables of form $I \otimes A$, the Kronecker product of a unit matrix in the ground's operator algebra with some operator A in the conductor's algebra: The coherence of terms of different charge matters only for experiments that do *not neglect the state of the ground that much.*

Whenever superselection rules seem to hold, suspect oversensitivity to the uncouthness of Bohr's complementarity, and seek to widen the empirical domain so as to bypass the rules in the wider domain. The rules may yet hold in the narrower domain, and physical laws may even be adequately expressible in relation to the narrower domain, so the rules can represent predilection rather than outright error. For example, a rule that all states are electrically neutral is stronger than superselection of charge. Yet almost all of physics may be fit into this straitjacket by compensating for any charge by a very distant shell of opposite charge. Charged states of the universe cannot be so treated, yet I perversely prefer the assumption of neutrality for the universe (if the universe is defined) to any less global superselection rule!

If the importance of coherence is difficult to contemplate, let "charge" signify momentum, and the complementary property, "position" or else the definability of well-localized electron orbitals. Correlated momentum superpositions are with respect to a momentum reservoir or "ground", in a 1,2 system isolated at definite total momentum. Subsequent study of such non-plane-wave orbitals in which the coherence of the plane-wave components matters, must make deeper reference to the "ground for momentum" than is possible through a reduced density matrix.

My earliest clear example (Lubkin, 1970) is where the "momentum" eigenstate 1,2 is an unexcited hydrogen atom in a large box (so that "plane wave" is finitized) and it is position of the electron, system 1, that is under consideration. The "laboratory" 2 is the proton. In respect of 2's fiduciae, the system 1 is localized to within ~ 1 Å. To the outer laboratory 3, the electron is as delocalized as the whole hydrogen atom, since 3's momentum ground is not the proton.

So even for position, momentum, and angular momentum, there is a poorly analyzed complex of *definition relative to fiduciae* beyond the classical relativity of translation and rotation in space–time (Lubkin, 1977, Mirman, 1979), classical shifts of fiduciae being powerless to change sharpness. There is yet an ill-defined story about observers untranslated and unrotated, but differing in details of what is "in" the system and what while *not "in" the system is yet not "out" enough for Landau tracing.*

[4]Von Neumann (1955, Section VI.2, footnote 212) credits Landau (1927) with discovery of the production of mixed subsystem states by partial tracing.

Caveat: Prior Correlations? Not only does correlation between laboratory and observed system develop through a single measurement but earlier correlations may exist from the measurements which prepared the state. The tensor-product representation $\Sigma a_i x_i y_o$ of the coming upon a state $\Sigma a_i x_i$ by a ready-to-measure device y_o *prior* to measurement of x by y in von Neumann's discussion of the quantum coherence developed by measurement may be faulted in perhaps being insufficiently coherent *before* the measurement. The product state is nevertheless enough for ad absurdum refutation of absolutely neglecting all "macroscopic coherences."

6.4 Connection between Smallness and Interference, Concluded. We have seen that the ontological complexity of quantum mechanics may in principle be visited upon entities within a grossly engineered "jelly" or "raster," without any regard to questions of atomic structure of the raster: $\hbar$ was not mentioned. These ontologies "internal" to the raster appear to be disparate from our own reality.

But even if our own reality as observed is narrowly considered in conventional particle physics, its status is muddied by questions of definition of important laboratory correlates, "grounds," which are not officially "in" the system, yet are not entirely to be neglected either.

What we interact with, our sequence of actual experiences, forms our notions of reality—a truism somehow tied to background reservoirs. The presumed 3-space context of traditional physics has to do with the rigidity of solids (Einstein, 1953; Lubkin, 1964). These things are not absolute. Much of their flux escapes personal choice, yet may have been chosen as part of Darwinian evolution of solid-associated life from prebiotic forms (Schrödinger, 1945).

If reality is a free-wheeling affair not yet pinned down even in particle physics, then it is presumptuous to say that a progression of choices of what is to be strongly interacted with may not induce an ontological drift, to a reality whose simplest entities are even more vaguely related to the atoms than is the grounded conductor.

The interference theorem limits this, yet one must not be overhasty. Is interference between economic states x and y, grossly defined entities, rejected because some detail different in x and y is not included in subsequent tests? Only if the detail may be isolated as a tensor factor within the economic model's framework of states, etc., with care to beware of rashly "tracing out" over associated "grounds" where inappropriate. Maybe the model misses the detail carelessly, and it can be put in as an inessential change yet so as to validate the interference theorem—but perhaps such management of a detail *is* a way to miss the point, to make an essential change.

Without deciding whether our reality can really drift so far away from atomic physics as to get tangled up in something like the "raster" within some new purely theoretical but decisive "ontological dynamics," we can try by experiment to see what interference effects there are, and therefore be guided by the real facts of reality in formulating how dirty our own quantum ontology can get.

7. QUANTUM LOGICS WITHOUT OUR REALITY BRANCHING

Relativity of reality is the most liberating notion in quantum mechanics, and a true quantum structure will impose it upon us. Interference of states that include us as observers is part of understanding the relationship of amplitudes to probabilities (Lubkin, 1979).

It is possible, instead, to imagine a quantum format analysis of an empirical domain, so limited as to preclude the inclusion of most systems, in particular not of us. The probabilities might even be known to appear owing to crudity in the techniques, like nonquantum coin tossing, while not being comprehended, however, as such within the given empirical domain.

If the domain described is closed in itself and does not reach out for the universe, it may, in being separate from issues of relative reality and presumed effects of the interference theorem, involve data fits with highly noncommutative small matrices.

This could be exemplified by an economic study divorced from the idea of the economist as a firm. The fact that one is free to instead consider larger systems that *do* comprise the observer, I consider more interesting than the notion of such new isolated quantum formats.

The idea of the "classical" jelly or raster, Section 6.2, may illustrate more clearly the possibility of an isolated quantum logic containing its own structure of relative reality for a system of internal observer states, but divorced from us. That was explained in terms of analogy between the classical equations of motion of small vibration and the Schrödinger equation, not in terms of an empiricism of states and tests. But we can tack one on. If the states are certain "bowing" procedures, the tests assignation of positive numbers $p_1,\ldots,p_b$ according to the response of b-channel acoustic analyzers, rather than probabilities, it is possible nevertheless to imagine MF as the tool for contacting the system.

Construction: Dials to Probabilities. Now, let the dials reading $p_1,\ldots,p_b$ be replaced by a Rube Goldberg apparatus: Small amounts of a dye are automatically introduced into b identical preparations in the proportions

p_1: ...:p_b, and a single photon is reflected back and forth through all the preparations. Its eventual absorption by the dye is considered certain, and leaves a signature in the appropriate vessel. This also escalates to ring a bell, so we know when the photon has been registered. We are sensitive to the exigencies of quantum logic, and are unaware that the whole p_1: ...:p_b information is potentially available at once nonstatistically on dials, and so reprepare everything after each single photon count. This clumsy "photon dial" introduces normalization of the $p_1,\ldots,p_b$, the photon registration probabilities being not p_i but $p_i/\Sigma_{j=1}^{b}p_j$.

However, we produce an analysis of an empirical domain by MF, the one-dimensional projections lead back to a Hilbert space. An analysis of the whole into parts according to tensor factorization and other "luxuries" of Section 10.5 could approach the nature of systems effectively from a standpoint of their own internal ontologies of convenience, even if our own ontology is not essentially involved.

8. LOGICS

Metaphysical Definitions. Domain, State, Test, Trials. An *empirical domain* is defined by a list of experimental *procedures*. I will require throughout that each procedure be "factorizable" into a procedure of *preparation*, and the execution of a *test*. The *i*th procedure of preparation will be said to prepare the *i*th *state*, A *test* or *b test* is clearly defined only if it has *b* specific possible outcomes, or "bins", numbered $1,\ldots,b$. A single *trial* of the test is a single performance of the procedure for preparing some state coupled with the procedure for effecting the test. The bins must be so defined that precisely one of the *b* possible outcomes actually happens, in any one trial. The process by which preparation of the *i*th state is coupled to executing the *j*th test to give a trial of *type* (i, j), must be empirically defined. Unless a method of coupling is procedurally given, the empirical domain has not been defined.

The experimenter determines p_{ijk}, the probability that a trial of type (i,j) eventuate in bin k out of the b_j possible bins of test j. This is done by rerunning the whole procedure for preparing state i and effecting test j, once for each single count in some bin, until very many counts, n_k in bin k, have been accumulated. Then $p_{ijk} = n_k(n_1 + \cdots + n_{b_j})^{-1}$ are the frequency-based empirical probabilities, subject to vicissitudes of statistical inaccuracy which will be overlooked: The "law of large numbers" and a *guarantee of independence of the trials* will be taken to justify that the p_{ijk} can eventually be found by experiment.

Definitions: Logic, State, Test. A *logic* is a p_{ijk} function, i.e., a map p: $(i,j,k) \to p_{ijk}$ from triples of labels, conveniently integers, to nonnegative numbers p_{ijk}. There are s state labels $i = 1, \ldots, s$, and t test labels, $j = 1, \ldots, t$. The map p need not be defined for all st values of (i, j). However, for such (i, j) that p is defined, the set of k for which it is defined is $1, \ldots, b_j$, the map $j \to b_j$ specifying the bin number of test j. *Extension and restriction of a logic* is defined by extension and restriction of its p function. Restriction of the logic function to a fixed state label i defines *state* i, restriction instead to a fixed test label j defines *test* j.

Remark: Histograms. A trial of type (i, j) is a procedure eventuating in a selection of one among b_j bins. Any histogram may also be thought of in this way. Unfortunately, an isolated histogram provides a definition of p_{ijk} only for one fixed type (i, j). Such information is a "logic" insufficient to yield separate pictures of a *plural* set of states i and a *plural* set of tests j. The separation of experiences (i, j) into a somewhat "objective" world of alternative states or inputs i, and tests or output procedures j, is what MF and such are about. Even if for each i of several, p_{ijk} is defined only for a unique j, the notion of state-test factorization is empty. In the sequel, we instead require many different tests, distinct j, for most single states i, and many states i available to most single tests j (Stapp, 1971, 1972).

9. THE MATRIX FORMAT, MF

This section states MF prescriptively; discussion follows in Sections 10–14.

9.1. Prescription, *b*-plex. The object is, given a logic, to associate a square $n \times n$ Hermitian nonnegative matrix P_i to each state i and another such matrix A_{jk} to each bin k of each test j, $(j,k) \to A_{jk}$. The value of n is unknown. However, the larger n is, the more parameters the given data p_{ijk} must determine. In the general spirit of a fitting of parameters one seeks a fit for the least n possible.

It is possible that distinct procedures of preparation i lead to identical state matrices P_i. The *state matrices* P_i are the familiar *density matrices* and have trace 1.

The test, bin matrices A_{jk} are the *acceptors*. Like the state matrices, these are unknown originally, and different bins of the same or of different tests may or may not eventually turn out to have equal acceptors. The acceptors need not have trace 1. Instead, the b matrices $A_1, \ldots, A_b$ for all the b bins of any one b-bin test or *b-test* are required to sum to the unit matrix, I. Such a list of nonnegative Hermitian matrices which sum to I will be called a *b-plex*.

The logic limits the matrices, by $\operatorname{Tr} P_i A_{jk} = p_{ijk}$.

MF, in summary, is a problem to find matrices $P_1, \ldots, P_s$; $A_{11}, \ldots, A_{1b_1}; \ldots; A_{t1}, \ldots, A_{tb_t}$, for the s states and t tests, the jth test having b_j bins, all $n \times n$ nonnegative Hermitian matrices,

$$P_i \geqslant 0, \qquad A_{jk} \geqslant 0$$

such that

$$\operatorname{Tr} P_i = 1$$

$$\sum_{k=1}^{b_j} A_{jk} = I$$

and

$$\operatorname{Tr} P_i A_{jk} = p_{ijk}$$

A_{jk} is a whole matrix, not the (j,k)th element of matrix A.

9.2. Generality, Classical Format, CF; Good for Ordinary Quantum Mechanics. Lubkin (1979) reviews how a Hermitian matrix A is associated to a traditional test procedure, so as to give $\operatorname{Tr} PA$ for the expectation value $\sum_{i=1}^{b} p_i a_i$ over the b bins, in a Schrödinger-cat-like context; but the reader will of course already be familiar with the $\operatorname{Tr} PA$ formula in quantum mechanics. Section 10 builds from this to MF, which encompasses ordinary quantum mechanics in the ordinary way, except for having n finite.

Significantly, MF encompasses classical logic. A classical MF solution is one with commuting matrices (see below), and conversely any finite classical logic can be expressed that way.

Perhaps there is a smallness of something *physical* that is responsible for the linearity that makes matrices relevant. Such smallness is likely to be challenged first in new physics, not in psychology, but who knows?

More General Formats? Real Format, RF. Logics, Section 8, have no evident linear structure; MF, Section 9, is committed to complex linearity. If the (state, A)→probability mapping is *required* linear in A, we get the $\operatorname{Tr} PA$ formula and MF. For literature seeking to deduce linearity, see Gleason (1953), Mackey (1963, Appendix). Measurement theory of a slightly nonlinear Schrödinger equation, perhaps through raster dynamics, might teach us something. For the present, I comment on the variants obtained by replacing complexes with reals or quaternions (Finkelstein, Jauch, Schiminovich, and Speiser, 1962).

Imitating MF with matrices of real numbers, *real format* or *RF*, can obviously be considered a special case, not a generalization of MF, but see the section on Nonconjugate Cousins below; also replace $U(n)$ by $O(n)$, and that's that for real quantum mechanics.

One might expect a similar comment for quaternionic quantum mechanics because of the representation of quaternion $d+ai+bj+ck$ as complex matrix

$$\begin{bmatrix} d+c(-1)^{1/2}, & a(-1)^{1/2}-b \\ a(-1)^{1/2}+b, & d-c(-1)^{1/2} \end{bmatrix}$$

However, the positive definite inner product $(q',q)=\Sigma_{i=1}^{n}\bar{q}_i'q_i$ of quaternionic n-tuples fails to square to form $\operatorname{Tr}P'(q')P(q)$ owing to noncommutativity of quaternions, making it doubtful that $2n\times 2n$ MF's $\operatorname{Tr}PA$ is relevant; more physically, putting simple systems together into compound systems is there a difficulty (Lubkin, 1979). Compounding, important itself, is also part of forcing the probabilities out from the mechanics. So though the abstraction of a lattice leads to a lattice of quaternionic subspaces, there may not be a true quaternionic quantum mechanics to worry about.

David Finkelstein points out that there are quantum logics with a finite number of elements. Those obtainable by imposing phase restrictions on state vectors in ordinary quantum mechanics will also fall, albeit awkwardly, within the scope of MF.

Classical Logic as a Special Case, CF. MF restricted to $n\times n$ diagonal matrices will be called *classical format*, CF, or CF_n. The rest of this section should make the sense of that clear. CF is later used to pit classical logic in CF_{n^2} form against $n\times n$ MF quantum logic, Section 15.2.

Classical logic has little to do with classical mechanics, more to do with the classical mystic attitude. Indeed suppose that every state is negligibly modified by every test in the empirical domain, so as to be available unmodified for other tests. Then one can replace the tests originally in the domain by one universal test, consisting of the succession of all the original tests, each one, furthermore, performed very many times. This universal test can in one such giant trial pick up the probability distributions over the bins of each original test. Pick some not too large N, and invent a finite list of new bins for the universal test which roughly designate these distributions, e.g., as follows. A new bin is specified by the list $(j,k)\rightarrow m_{jk}/N$, where the m_{jk} are positive integers $\leqslant N$ such that

$(m_{jk}-1)/N<p_{jk}\leqslant m_{jk}/N$, and p_{jk} is the known but formally unknown probability to be remeasured in a new trial that the state being universally tested drop in bin k of original test j. We expect only one bin of the universal test to register for every single universal trial.

Notice the disrespect for the *original* notion of single trial. The idea that the state is negligibly disturbed by the original tests has here been made responsible for this disrespect. Instead, the mystic attitude, a conviction of apprehensibility of a concrete real world, may be made responsible: Even if tests spoil the state, necessitating fresh copies for plural trials, this is regarded as incidental; the refreshment technique is considered a clumsy substitute for better ideal tests which would gain equivalent information without disturbing the state. The mystic notion of the concretely real state existing by itself with properties independent of any process of observation amounts to this, our learning properties without essentially spoiling the isolation.

Indeed, this "mystifying" device of many trials allows the association in quantum mechanics of a state matrix P to each state, mod conjugation, if there are enough types (i, j) of experiment. In quantum mechanics with completion of states by "rouletting" or mixing, the states are mixtures of pure states; let these mixtures be called generalized wave functions. {The map $i\rightarrow[(j,k)\rightarrow p_{ijk}]$ is a generalized wave function appropriate for logics in general.} If the multiplicity of trials required to fix a generalized wave function is disregarded by not distinguishing multiple preparations of the same state from a single preparation, then all logics look classical. Quantum logic is existentialist in insisting on the priority of the individual experience over any other notion, refusing to bury the elemental trial behind a structure of repetitions before presenting it as a "trial" to the apparatus for generating theoretical elements to represent the states and tests. The repetitions are fine, but the elemental state, test, and trial without repetitions, are explicitly not rejected as part of the empirical domain.

The "mystifying" transformation of any logic produces an empirical domain and a logic where the values p_{ijk} are all 0 or 1, and there is only one, universal test. The index j is not needed; p_{ik} is the probability that state i register in bin k of the *one* test. Also, $p_{ik}=1$ for precisely one k for each i. Identify states which yield the same probabilities for all tests, i.e. here for our one test. Then at most one state i has $p_{ik}=1$ for given k. Drop bins k where no state ever registers. Renumber the remaining b bins so that $p_{ik}=\delta_{ik}$.

There is now a set of b elements, the states, and one test, where state k simply proves itself on one trial by activating bin k. The logic is degenerated to an Aristotelian labeling of distinct things by distinct names. The classical lattice of sharp questions (Section 10.3) corresponds to 2-tests formed from subsets S of the b-element set of bins, by lumping registration in any bin belonging to S into "yes," registration not within S into "no."

A fit in MF with $n \times n$ matrices which commute is classical, with an n-bin universal test if there is a repertory (Section 11). $U \cdot U^{-1}$ diagonalize all the matrices simultaneously. $\mathrm{Tr}\, PA$ becomes $\Sigma_{i=1}^{n} P_i A_i$ in terms of the diagonal elements P_i and A_i. $\mathrm{Tr}\, P = 1$ becomes $\Sigma_{i=1}^{n} P_i = 1$. Nonnegativity becomes nonnegativity of the diagonal elements. Each state P is a convex combination of the n pure states where a single diagonal element is 1. The sharp questions $(E, I - E)$ of Section 10.1 are given by stating which subset of the diagonal corresponds to the "1's" of E. The n one-dimensional projections are the n acceptors of the universal test.

If a repertory is unavailable, the convex body of available states may be smaller than the simplex generated by the n one-dimensional projections above, and there may also be a restriction of the convex body of questions. A commutative fit is classical nevertheless in the sense (von Neumann, 1955) that empirically unattainable pure states and sharp questions may be postulated within n by n MF without enlarging n, where the $\mathrm{Tr}\, PA$ taken between the pure states and the acceptors of the sharp questions are all 0 or 1. All nontrivial probabilities look like those generated by convex combination in the manner of "rouletting" (Section 10).

Bloc Format, BF. CF is an extreme case of $\mathrm{BF}(n_1, \ldots, n_B)$, MF restricted to matrices in diagonal block form, nonzero elements in n_1 by $n_1, \ldots, n_B$ by n_B nonoverlapping diagonal blocs, $n_1 + \cdots + n_B = n$, namely, $\mathrm{CF} = \mathrm{BF}(1, \ldots, 1)$. BF is equivalent to superselection rules: superobservables commute, hence can be simultaneously diagonalized, leading to BF, and conversely the BF bloc projectors are superobservables.

Nonconjugate Cousins. In BF, hence CF, and in RF, states and tests are specialized from MF in having zeros in certain places. If the zeros condition is lifted from either state matrices alone, or from test matrices alone, the $\mathrm{Tr}\, PA$'s are unchanged, hence the logic is invariant to such partial format repeal. If there are enough states P, for example, a unitary or antiunitary conjugation leaving all P fixed would be the identity, and so such partial repeal (here for the tests) goes beyond ordinary conjugation. Thus an MF computation on a BF-, CF-, or RF-compatible logic is likely to find some nonconjugate cousin, not simply a conjugate, of a BF, CF, RF solution. Specializations will be seen most clearly in computations restricted a priori. This comment neglects the inequalities of nonnegativity of MF, which if there are enough states and tests *can* lock the arbitrariness down to conjugations: Section 11.

9.3. The Size *n* of the Matrices. The larger n, the more unknown parameters are available to fit the data with $n \times n$ MF. Precisely how many MF logics are there, for given n? That is, determine the real dimension $d(n)$ of the manifold of functions $(i,j,k) \to p_{ijk}$ generated by $\operatorname{Tr} P_i A_{jk} = p_{ijk}$.

Theorem 2: Matrix Logic Dimension Theorem.

$$d(n) = (n^2 - 1)(s-1) + n^2 \sum_{j=1}^{t} (b_j - 1)$$

Proof. Of the relevant matrices I; $P_1, \ldots, P_s$; $A_{11}, \ldots, A_{1b_1}$; $\ldots$; $A_{t1}, \ldots, A_{tb_t}$, I is known, the state matrices P_i each provide $n^2 - 1$ unknowns, not n^2, as their traces are known; and of the b_j acceptors of test j, $b_j - 1$ of them completely fix the remaining one, and have n^2 unknowns apiece; nonnegativity does not figure in a dimension count. This so far provides a count of $(n^2 - 1)s + n^2 \sum_{j=1}^{t} (b_j - 1)$ unknowns for the real dimensionality of the family of lists of matrices which fulfill MF for arbitrary p_{ijk}. From this must be subtracted $n^2 - 1$, the dimensionality of $SU(n)$, because $SU(n)$ mod its center is one of the two components of the least (Section 11 or Appendix C) effective group of transformations of the matrices under which a logic remains invariant. ■

Remark. Dimension Jump.

$$d(n+1) - d(n) = (2n+1)m$$

where

$$m = s - 1 + \sum_{j=1}^{t} (b_j - 1)$$

measures the increase in the dimensionality of the logics made available by stepping $n \to n+1$. The number of independent states and bins is $m+1$; m is independent of n.

Given the problem of fitting empirically determined p_{ijk}, a "solution" can likely be found by taking n sufficiently large. For example, if $n^2 \geqslant m + 2$, the condition that all $m+2$ matrices (counting I) considered as vectors (Section 12) lie in a common n^2-dimensional vector space, becomes vacuous. Does this indicate that after all "quantum logic" is tautologic, or vacuous?

The situation in physics indicates otherwise.

The most satisfactory sort of solution would be rigid or unique, in the sense that only conjugations in MF or orthogonal transformations in the DF's (Section 12) would preserve the fit. However, the jump by the large number $(2n+1)m$ of parameters available for fitting data is most likely to lead from some n where no good fit has been found, to the next size $n+1$, where a fit not only is possible but also is highly nonrigid. So a good rigid fit is implausible, as an accident. Rigidity in itself provides some significance.

A nonfortuitous quantum logic fit may also have some nonrigidity, perhaps associated with superselection rules, if there is a deficiency of states and tests. But as m is enlarged within the same qualitative "domain of empiricism," a true quantum logic would rigidify, except possibly for superselection rules not transcended by the empiricism, in its proper n, and would stay reasonable in the same n, while new techniques increase m beyond where a fortuitous fit is plausible.

For much larger n, a classically logical fit should be possible, comprised within MF (Section 10.2). The dimensionality $d_{\text{classical}}(n)$ of the classical logics within $n \times n$ MF is easily calculated by requiring that the matrices be diagonal: CF. Then a density matrix provides $n-1$ free parameters, an acceptor n free parameters. There are now

$$d_{\text{classical}}(n) = (n-1)s + n \sum_{j=1}^{t} (b_j - 1)$$

free parameters to fit the empirical probabilities, with no effective alteration of the parameters leaving the logic unchanged.

The likelihood of fitting measured p_{ijk} by *very many* parameters should make some CF fit inevitable. So if quantum logic is not tried, it may not be missed.

10. STRUCTURE OF THE TESTS

Because of the formal presentation of MF in MB (Lubkin, 1974a, b), I feel free to discuss the sense of that work here more flexibly.

10.1. Rouletting vs. the Register Method. My *states* are the well-known density matrices of Landau (1927), von Neumann (1927, also 1955, Section IV.2, footnote 172), and Weyl (1928), but my *tests* are somewhat novel.[5] To motivate this new look at tests, recall how wave functions were generalized to density matrices. Indeed, this happened in *two* distinct ways.

[5] R. Giles (1970) independently generalized the 2-tests.

The *first* way, which I call *rouletting* of states, came about because it was desired that a procedure in which a random process like spinning a roulette wheel begins the procedure, with the outcome of this random process then designating which of several possible subsequent state preparations will ensue, should also be regarded in its entirety as a single legitimate procedure for preparing a state. This requirement, that the set of states be complete under rouletting, amounts to convex completion of the logics p_{ijk} on the state index i. Of course the main historical motive for this completion was the incorporation of statistical ensembles into quantum mechanics; but an incidental advantage is that procedures can be legal preparations of states even before one can argue which among them prepare pure states.

The most evident parallel lesson concerning the tests here would be to study the rouletting of test procedures, i.e., of the logics on the test index j, but this does not quite lead to MF, which is simpler and more general.

The *second* way in which wave functions lead to density matrices comes from the desire to focus attention on a *subsystem*: The density matrix $P_{i_1i_2,\, j_1j_2}$ for a compound system is replaced by $\rho_{i_1,\, j_1} = \Sigma_{i_2=j_2} P_{i_1i_2,\, j_1j_2}$. The partial trace performed over the 2-associated indices represents the focusing of attention upon system 1. This is the *Landau tracing* already mentioned in Section 6.3, a topic I find fascinating (Lubkin, 1978, 1979). Even when P is a pure state, the subsystem's representative matrix ρ is usually a mixture. Thus, in quantum mechanics, the intent to ignore any part of a state has the power of mimicking the straightforward introduction of probabilities by means of rouletting.

It is by consideration of a multisystem complex in analogy to the system 1, system 2 context of Landau tracing that I arrive at MF by what I will call the *register method*.

In conventional quantum mechanics, the state P is confronted by some observable, a normal or Hermitian operator Y, with outcomes corresponding to distinct eigenvalues $y_1, \ldots, y_b$ of Y (Wigner 1952, Lubkin 1979). The probability p_k of the kth outcome y_k is $p_k = \mathrm{Tr}\, PE_k$, where E_k is Hermitian projection on the eigenspace $y_k(Y)$ of Y corresponding to eigenvalue y_k. $E = (E_1, \ldots, E_b)$ is a b-plex (Section 9.1) of special form, its coordinates E_k being mutually orthogonal projections. I call such b-plexes *sharp*. Because the conventional Dirac (1947) or von Neumann (1955) observables correspond to sharp b-plexes, I call them *sharp tests*.

The generalization of these conventional sharp tests by the register method goes as follows. System 1 is the system of interest associated with n-dimensional Hilbert space, prepared in some state with density matrix ρ. System 2, the *register*, is associated with a b-dimensional Hilbert space, and is prepared in an initial pure state e_1. The overall initial state is then

$P=\rho\otimes e_1$. Then the system 1,2 complex, a system whose state is associated with the nb-dimensional tensor-product Hilbert space, is made to undergo some unitary motion U, leading to new overall state $UPU^{-1}=U\rho\otimes e_1U^{-1}$. Finally, the observer collects an outcome by observing the register, asking impulsively in which of the orthogonal register pure states $e_1,\ldots,e_b$ it is. Thus, the experiment has b possible outcomes, and the probability of the kth outcome is $p_k=\mathrm{Tr}\,U\rho\otimes e_1U^{-1}1\otimes e_k$. Were we to regard the whole story up to developing UPU^{-1} as the preparation of a state *of the system-register complex*, this queried by the sharp final b-test with b-plex $(1\otimes e_1,\ldots,1\otimes e_b)$, we would evidently *not* be leaving the domain of ordinary, sharp observables. We wish instead to regard this as the preparation of *state* ρ *of the system of interest alone*, followed by subsequent procedures which are all to be reckoned as part of a test with b outcomes also performed upon the system of interest alone, in that the probabilities are to be calculated by means of a formula of form $p_k=\mathrm{Tr}\,\rho a_k$ from a b-plex a of b acceptors $a_1,\ldots,a_b$, Hermitian operators or $n\times n$ matrices defined *over system*-1 *Hilbert space*. The coupling with the register and the $U\cdot U^{-1}$ motion are to be reckoned as part of the test, not the state, and the register is laboratory equipment outside of the Hilbert-space story. b-plex a depends on the motion U, and in such a way that, as U runs over all nb by nb unitaries, a runs over *all* $n\times n$ b plexes. (The verification is in MB.)

Because this register method of doing an experiment can reach any b-test, it follows that the traditional observables, whose b-plexes are sharp, correspond to a restricted class of tests: such as may be considered determinations of eigenvalues of observables. Hence, in an investigation wherein the states and tests are explained as procedures not known to be related to the Dirac or von Neumann observables of a mechanical theory, it would be presumptuous to require that the probabilities be fitted with sharp b-plexes for b-tests, rather than by the more general b-plexes.

Let us return to rouletting tests: If we start with several sharp b-tests, with sharp b-plex $E^r=(E_1^r,\ldots,E_b^r)$ for b-test r, then the rouletted b-test performed by selecting b-test r with probability p_r will have b-plex $A=(A_1,\ldots,A_b)$ with $A_k=\Sigma p_rE_k^r$, and will not usually be sharp. We may write $A=\Sigma p_rE^r$. A is a *convex combination* of the E^r. For example, let a conventional atomic system have energy values $h_1,h_2,h_3,\ldots$ and total angular momentum values $j_1,j_2,j_3,\ldots$, and let the experiment be, to toss a coin, and measure energy if the coin comes up heads but angular momentum if tails, and nevertheless to simply record outcome k if either, in the case energy is measured, the value obtained is h_k, or if angular momentum is measured, the value obtained is j_k: here *two* sharp observables are being convexly combined. Such rouletting shows us, perhaps more easily than the register method, that the observable of conventional quantum mechanics should be generalized. The register method is, however, more powerful, because rouletting fails to generate the set of *all* b-plexes from the sharp ones, when $b \geqslant 3$ (see MB). Those b-plexes that may be generated by convex combination of the sharp ones, I call *undersharp*.

The rouletting together of sharp tests is a test analog of the rouletting or convex combination or mixture of states, and the register method is a test analog of Landau tracing—because as in Landau tracing, there is a system 2 with subordinate status, the state ρ and the b-test a referring only to the $n \times n$ system 1 of interest. The fact that the b-plexes representing b-tests with $b \geqslant 3$ attained by the register method are *more general* than the undersharp b-plexes attained by rouletting sharp b-tests, is a fact about tests that stands in contrast to the fact about states, namely, the development of the *same* state density matrices by Landau tracing from an enveloping pure state as by the rouletting of pure states of the system of interest.

Are the b-plexes a sufficient generalization of the traditional observables? It is easy to see that this is so (MB) if the probabilities of outcomes are to depend linearly on the von Neumann density matrix of the state: For b-plex $A = (A_1, \ldots, A_b)$, the condition $\Sigma A_k = I$ is imposed by normalization of the probabilities, A_k Hermitian by their reality, $A_k \geqslant 0$ by their nonnegativity.

It might be expected—perhaps from a feeling for symmetry between past and future—that just as "in" states, our states, become density matrices upon convex completion, never mind b-plexes: tests considered as "out" states should also become density matrices. They do not. The lack of in, out symmetry is familiar in treating the unpolarized condition in nuclear physics: at the "in" end one *averages*, while at the "out" end one *sums*. This asymmetry is exploited in determining some spins from others without using polarization experiments, by comparing cross sections for inverse processes (e.g., Frazer, 1966). This is not to deny that outcomes of tests can help prepare states: e.g., see Section 10.4.

10.2. Convexity. The story of convex combination of sharp b-plexes, parallel to the laboratory procedure of rouletting sharp b-tests, of course need not be restricted to sharp tests; under unrestricted convex combination, the b-plexes of $n \times n$ matrices form a convex body for each b, n. In the norm topology wherein the norm of a b-plex is the sum of the absolute squares of all the matrix elements of all b matrices, the body is also *compact*. Hence it is the convex completion of its set of extreme points, and it is natural to seek information about the extreme points. It is easily shown (MB) that all sharp b-plexes are extreme. The undersharp b-plexes of course comprise the convex body generated from the sharp b-plexes (for each n); equivalently, the convex body generated by the commutative b-plexes (those whose b matrices commute), and the statement that (for $b \geqslant 3$) there are yet other b-plexes is equivalent to the existence of nonsharp extreme b-plexes. This is done in MB by determining all extreme b-plexes of 2 by 2 matrices. $b=1$ is of course the uninteresting degenerate case with only one 1-plex (I), necessarily extreme. For $b=2$, the extreme 2-plexes are the sharp 2-plexes (P,Q). Among these, those where $\operatorname{rank} P = \operatorname{rank} Q = 1$, specify the extreme 2-plexes with no zero matrices. b-plexes without zero matrices will be called *0-free*; the $(b+1)$-plexes, $(b+2)$-plexes,... obtained by listing extra zero matrices are all extreme if and only if the 0-free core obtained by omitting zero matrices is itself extreme. The *extraneous* extreme 0-free b-plexes of 2×2 matrices are the following 3-plexes and 4-plexes. Each acceptor is a 2×2 Hermitian matrix of form $|a|I + a\cdot\sigma$, where σ are the three Pauli matrices and a is a real 3-vector, hence the acceptor can be tagged in terms of its real 3-vector a. The extraneous extreme 0-free 3-plexes correspond to vectors a_1, a_2, a_3 which form a triangle of unit perimeter and nonzero area, and the extraneous extreme 0-free 4-plexes to a_1, a_2, a_3, a_4, which form a nonplanar quadrilateral of unit perimeter. These triangles and quadrilaterals are noncommutative. *There is no extraneous 2-plex for any n.* Construction of extraneous b-plexes for $b \geqslant 3$ for $n>2$ by building upon the $n=2$ example is easily done.

These results of MB were obtained in part with the help of two notions of reducibility, which I called *segmentation* and *1-elimination*. A b-plex irreducible in both senses would be *1-free*, namely, none of its acceptors would have 1 as an eigenvalue, and also no nontrivial sum of some of its acceptors would be a projection. The complete reduction of a general b-plex is obtained by first making a unitary conjugation of all the matrices so as to bring all 1 eigenvalues into one bloc, say at upper left, which bloc becomes a sharp direct factor; call the bloc projector E^0. The complementary lower-right bloc yields a unique finest segmentation, namely, a list of Hermitian projections $E^1,\dots,E^s$ (s segments, $s \geqslant 0$) such that some of the lower-right matrices add up to E^1, some others to $E^2,\dots$, the rest to E^s, with $E^1 + \cdots + E^s$ being the lower-right-bloc unit matrix, i.e., projection on lower-right space. Unitarily conjugate to diagonalize $E^1,\dots,E^s$, to complete the reduction. A list of b nonnegative Hermitians which sum to projection E is a *(b, E)-plex*. The complete reduction described yields a possible upper-left sharp (b, E^0)-plex, then a (b, E^1)-plex, a (b, E^2)-plex, etc., each latter irreducible in both senses, racked together as a direct sum. *The whole thing is extreme if and only if each irreducible ingredient b-plex is itself extreme.*

Both of the above notions of reducibility of MB are special cases of the more conventional (unitary) reducibility, wherein if a set of matrices is brought by unitary conjugation into bloc form, that counts as a reduction without further ado. Indeed the sum of the b matrices of each bloc comes out the bloc projection operator, with those for distinct blocs orthogonal, but unlike the more special notion of segmentability, it is not required that each whole original matrix contribute to only one segment bloc. I have unfortunately not been able to either prove or disprove the *conjecture* that this less specialized notion of reducibility also has the property that *a complete reduction is extreme if each irreducible component is extreme*: That the composite is not extreme if a component is not, is shown by using the convex combination which displays the nonextremity of the component given nonextreme, with the other components unvaried.

I have been able to shift the essence of the matter away from the topic of convexity to something more purely algebraic: Let E_V be orthogonal projection on subspace V of finite-dimensional Hilbert space $\mathcal{H}$. The set of operators of form $E_V M E_V$, as M runs over all linear operators on $\mathcal{H}$, is a vector space of dimension $(\dim V)^2$ which I call the *quadratic space* $Q(V)$ *of* V. A list of subspaces of $\mathcal{H}$, or the corresponding list of projections, is *quadratically independent* if the corresponding quadratic spaces are linearly independent. Then the truth of the following proposition would prove my above conjecture about reducibility: *Let* $E_1,\ldots,E_b$ *and* $F_1,\ldots,F_b$ *each be a quadratically independent list of projections. Let* $E_1+\cdots+E_b=E$, *let* $F_1+\cdots+F_b=F$, *and let* E,F *be quadratically independent. Then* (conjecture) $E_1+F_1,\ldots,E_b+F_b$ *is also quadratically independent.*

For a list of two items, the condition of quadratic independence is equivalent to linear independence, which is therefore implied for E,F; yet it is amusing that it can be worded all in terms of quadratic independence. Indeed, it can be shown that, although the quadratic *space* $Q(V)$ varies for fixed V as the inner product in $\mathcal{H}$ is varied via alternative nonsingular Hermitian forms, the quadratic *independence* of a list of spaces is invariant with respect to choice of inner product. Therefore, although the proposition appears to depend upon choice of inner product, it does not in fact do so if restated in terms of spaces rather than projections. It is really about vector spaces, not Hilbert spaces. Perhaps this generality will encourage interest.

10.3. Lattices and Logics. The lattice-theoretic approach to quantum logics (Birkhoff and von Neumann, 1936; Jauch, 1968; von Neumann, 1955, p. 247) deals with the sharp questions: A question or 2-plex (A,B), $A+B=I$, $A\geqslant 0$, $B\geqslant 0$ may be represented by its first matrix A, $0\leqslant A\leqslant I$. It is sharp if and only if A is a projection. The sharp questions' projections A are one–one with the subspaces of n-dimensional Hilbert space. These form a lattice under inclusion of subspaces, with set-theoretic intersection as meet and linear envelope of set theoretic union as join. Partial analogy with "and" and "or" in Boolean logic helps motivate the term "quantum logic." My use of "quantum logic" departs from this chiefly because my tests are more general than sharp questions.

Remark: A Metaphysical Advantage of Mixture (Lubkin, 1976). The lattice of sharp questions for the quantum logic of $n\times n$ matrices is disconnected into $n+1$ components: Questions $(A,I-A)$ has *dimension* $\dim 1(A)$ which ranges over $0,1,\ldots,n$ (and appropriately varying unitary conjugation does connect d-dimensional A to standard diagonal form). Similarly for the *sharp* b-tests even if $b>2$: $(A_1,\ldots,A_b)$ maps to a dimension-tuple component label $(d_1,\ldots,d_b)$ with $d_1+\cdots+d_b=n$ and $d_i=\dim A_i$.

Perhaps *physical* tests should not be so *formally* disconnected owing to the format of quantum logic alone. This worry may have spurred von Neumann's interest in continuous geometry free of a discrete dimension function (1962), aside from the issue of "modularity." However, the convex body of *all* b-plexes, b and n fixed, is of course connected! And it is impossible to decide without trials whether some test is (approximately) extreme; this quality emerges only in a calculated fit of many state-test trial probabilities, together. So interpolating from one sharp test to another of different dimension structure through mixtures physically bridges the dimension gaps, even if no *explicit* recourse to a roulette wheel is allowed. Ordinary Hilbert-space based physics is connected regardless of the discreteness of ordinary dimension.

10.4. Undersharpness. The main reason why the extraneous tests and mixed tests have been neglected is that it is enough to apply sharp tests, indeed sharp 2-tests, to define the states empirically: people have sought the *simplest* tests, indeed mainly those corresponding to Hermitian-observable conserved quantities, in order to get most directly at the states. The register method nevertheless shows how to achieve the most general tests. Between these extremes there are the undersharp tests, those which can be produced from the sharp tests by rouletting alone. More generally, what circumstances will guarantee undersharpness? I quote some undersharpness theorems from MB, then go more fully into some details omitted from that account.

The *fusion* of bins, that is the combining of counts from two or more outcomes originally reckoned as distinct into a single outcome, preserves undersharpness.

If the *sum* of the maximum eigenvalues of all but one of the acceptors of a b-plex is no greater than 1, then the b-plex is undersharp: crudely, a situation dominated by one *beam dump* bin is not a way to achieve an extraneous test. This provides a solid undersharp portion of the body of all b-plexes, hence the undersharp b-plexes and the general b-plexes of $n \times n$ matrices *share the same dimension*, $(b-1)n^2$.

So far, the tests could be destructive, a trial leaving only a count in a bin and no final state at all. Knowledge gained by applying such possibly destructive tests, that a certain procedure does indeed produce a particular "final" state P, is nevertheless useful in that the state-forming procedure can be carried out anew short of destruction, in order to provide P for some other purpose.

By a *filtration bin* I mean a test bin of this sort which also serves up a final state in the *usual* quantum-mechanical fashion: After normalized initial vector state x registers in a bin whose acceptor is projection E, the output state is proportional to Ex. The density matrix for this, unnormalized, is $|Ex\rangle\langle Ex| = P' = E|x\rangle\langle x|E = EPE$, where $|x\rangle\langle x| = P$ is the prior state. $\mathrm{Tr}\, EPE = \langle Ex|Ex\rangle$ is the probability of E-registering, so the reduced normalization is correct for a new "state" matrix P' which will yield the expected number of counts into a subsequent bin with acceptor A when substituted into the formula $N\mathrm{Tr}\, P'A$, where N is the number of trials prior to the E selection, if a "success" involves both E selection and subsequent A acceptance.

If instead of one projection E, there is a sharp b-test $(E_1, \ldots, E_b)$ composed entirely of E_i which in this way pass on the state, then the final counts into a subsequent bin with acceptor A is $N\mathrm{Tr}\, P'A$, where $P' = \Sigma_{i=1}^{b} E_i P E_i$, and $\mathrm{Tr}\, P' = 1$. This delicate sort of nondestructive sharp test $(E_1, \ldots, E_b)$ is a *filtration through all bins*, and the process $P \to P'$ on states induced by such a filtration is von Neumann's "process of measurement." Rouletting states of course generalizes the $\Sigma E_i P E_i$ formula (von Neumann, 1955) to mixed states P. Production of a correlated system–observer state $\Sigma a_i x_i y_o \to \Sigma a_i x_i y_i$ followed by Landau tracing over the observer space (Lubkin, 1979) produces the same result.

> *Theorem 3: Bifurcation Theorem.* If the bth bin of sharp b-test $(A_1, \ldots, A_b)$ outputs in the manner of a filtration, and if this bin is subjected to bifurcation by application of further filtration according to projections E, F, $EF = 0$, $E + F = I$, E leading to a new bth bin and F to a $(b+1)$st bin, then the resulting $(b+1)$-plex is $(A_1, \ldots, A_b E A_b, A_b F A_b)$, sharp if $A_b E = E A_b$, undersharp generally.

Proof. The output from the original bth bin is $P' = A_b P A_b$, if P is the original state, for subsequent use in computing the probabilities of final bth bin counts, $\mathrm{Tr}\, P'E$, and final $(b+1)$st bin counts, $\mathrm{Tr}\, P'F$. These probabilities are therefore $\mathrm{Tr}\, A_b P A_b E$ and $\mathrm{Tr}\, A_b P A_b F$. These can be cycled into form $\mathrm{Tr}PA'_b$, $\mathrm{Tr}PA'_{b+1}$, with $A'_b = A_bEA_b$, $A'_{b+1} = A_bFA_b$. The other acceptors are unchanged.

The $(b+1)$-plex thus established is sharp if $A_bE = EA_b$: It is easily verified that each $A_1, \ldots, A_{b-1}$, A_bEA_b, A_bFA_b squares to itself, that the product of any two is 0, and that they add to I.

It is a little more difficult to argue undersharpness when $A_bE \neq EA_b$. Consider the subspace Im A_b, A_b being a projection because the original b-plex is given sharp. In this subspace, (A_bEA_b, A_bFA_b) is a 2-plex, A_b being the appropriate unit matrix. Since every 2-plex is undersharp, there exist coefficients $p_i > 0$ with $\sum p_i = 1$ and sharp 2-plexes (E_i, F_i) with $E_iF_i = 0$, $E_i + F_i = A_b$, such that $(A_bEA_b, A_bFA_b) = (\sum_i p_i E_i, \sum_i p_i F_i)$. Then $(A_1, \ldots, A_{b-1}, A_bEA_b, A_bFA_b) = \sum_i p_i(A_1, \ldots, A_{b-1}, E_i, F_i)$, presents the bifurcated $(b+1)$-plex as a convex combination of sharp $(b+1)$-plexes. ∎

Theorem 4: Bifurcation of an Undersharp Test. Suppose several *sharp* b-tests, $(A_1^k, \ldots, A_b^k)$, each giving rise to a $(b+1)$-plex by means of (E, F) bifurcation as above of the bth bin, with the whole procedure rouletted with probabilities p^k: these are written with superscripts to emphasize that they are not the p_i of the last argument. Then the resulting $(b+1)$-plex is

$$\left(\sum p^k A_1^k, \ldots, \sum p^k A_{b-1}^k, \sum p^k A_b^k E A_b^k, \sum p^k A_b^k F A_b^k\right)$$

and is undersharp. If the rouletted b-plex prior to bifurcation is denoted $(A_1, \ldots, A_b)$, then this resulting $(b+1)$-plex *cannot* in general be computed from $A_1, \ldots, A_b, E, F$ alone, not even if A_b commutes with E; the rouletting itself must be given, as above.

Furthermore, *if E is 1-dimensional*, and if the undersharp b-plex is physically produced as a specific rouletting of sharp b-plexes, not only is the final $(b+1)$-plex determined, but also the manner in which it is rouletted from sharp $(b+1)$-plexes.

Proof. The resulting overall $(b+1)$-plex is obvious. It is undersharp, because it is a convex combination of $(b+1)$-plexes known to be undersharp from Theorem 3.

That the result cannot be expressed in terms of the $A_i = \sum_k p^k A_i^k$ and E, F alone is indicated by the difficulty of bringing both A_b^k together in $\sum_k p^k A_b^k E A_b^k$. Could this be done, say, if $EA_b^k = A_b^k E$, then the expression would collapse to A_bE (and the F term would collapse to A_bF). But if only $EA_b = A_bE$ is given, commutativity with E is not known for each A_b^k separately.

To prove that this is not merely algebraic clumsiness but that the p^k, A_b^k must indeed be separately given, an example is presented wherein two different roulettings which produce the same undersharp 2-plex of 2×2 matrices though subjected to the same (E,F) bifurcation yet produce *distinct* 3-plexes.

The 2-plex is

$$\left[\begin{bmatrix} \frac{1}{2} & 0 \\ 0 & \frac{1}{2} \end{bmatrix}, \begin{bmatrix} \frac{1}{2} & 0 \\ 0 & \frac{1}{2} \end{bmatrix}\right]$$

Both bins happen to have the same acceptor. Bin 2 is to be bifurcated into new bins 2 and 3, using

$$E=\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, F=\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$$

and two different roulettings.

The first rouletting is

$$\left[\begin{bmatrix} \frac{1}{2} & 0 \\ 0 & \frac{1}{2} \end{bmatrix}, \begin{bmatrix} \frac{1}{2} & 0 \\ 0 & \frac{1}{2} \end{bmatrix}\right] = \frac{1}{2}\left(\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}\right) + \frac{1}{2}\left(\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}\right)$$

The new acceptor for bin 2 is

$$\frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} \frac{1}{2} & 0 \\ 0 & 0 \end{pmatrix}$$

Thus the final 3-plex is

$$\left[\begin{bmatrix} \frac{1}{2} & 0 \\ 0 & \frac{1}{2} \end{bmatrix}, \begin{pmatrix} \frac{1}{2} & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & \frac{1}{2} \end{pmatrix}\right]$$

The second rouletting is

$$\left[\begin{bmatrix} \frac{1}{2} & 0 \\ 0 & \frac{1}{2} \end{bmatrix}, \begin{bmatrix} \frac{1}{2} & 0 \\ 0 & \frac{1}{2} \end{bmatrix}\right] = \frac{1}{2}\left[\begin{bmatrix} \frac{1}{2} & -\frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} \end{bmatrix}, \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix}\right] + \frac{1}{2}\left[\begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix}, \begin{bmatrix} \frac{1}{2} & -\frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} \end{bmatrix}\right]$$

The new acceptor for bin 2 is

$$\frac{1}{2}\begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix}\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}\begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix}$$

$$+\frac{1}{2}\begin{bmatrix} \frac{1}{2} & -\frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} \end{bmatrix}\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}\begin{bmatrix} \frac{1}{2} & -\frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} \end{bmatrix}=\begin{bmatrix} \frac{1}{4} & 0 \\ 0 & \frac{1}{4} \end{bmatrix}$$

This second rouletting gives final 3-plex

$$\left(\begin{bmatrix} \frac{1}{2} & 0 \\ 0 & \frac{1}{2} \end{bmatrix},\begin{bmatrix} \frac{1}{4} & 0 \\ 0 & \frac{1}{4} \end{bmatrix},\begin{bmatrix} \frac{1}{4} & 0 \\ 0 & \frac{1}{4} \end{bmatrix}\right)$$

distinct from the result from the first rouletting.

Observe that in this example, counter to the suggestion that bifurcation be independent of the rouletting, the acceptor A_b of the bin split *commutes* with E.

Finally, the inductive succession of roulettings *when E is one dimensional*, is argued: A_bEA_b, being of rank no greater than E, is itself proportional to a one-dimensional projection E': $A_bEA_b=pE'$, $0\leqslant p\leqslant 1$. Then, $(A_bEA_b,A_bFA_b)=(pE',A_b-pE')$. This is to be written as a convex combination of 2-plexes (E_i,F_i) with $E_i+F_i=A_b$, and coefficients of convex combination $p_i>0$, $\Sigma p_i=1$. In particular, $\Sigma p_iE_i=pE'$. This puts pE' internal to the simplex generated by the E_i. But pE' lies on an extreme ray (Appendix B), hence is internal only to one-dimensional simplexes. Hence each E_i is either E' or 0: we are past worrying about noncommutativity. Thus, $\Sigma p_iE_i=(p_1+\cdots+p_r)E'+(q_1+\cdots+q_z)0$ with $p_1+\cdots+p_r=p$ and $q_1+\cdots+q_z=1-p$. The rouletting coefficients are $p_1,\ldots,p_r$, $q_1,\ldots,q_z$, the p's being associated with sharp question (E',A_b-E') and the q's with sharp question $(0,A_b)$.

In Theorem 4, the resulting $(b+1)$-plex is independent of the rouletting prior to bifurcation, to the extent that if the $A_b^kEA_b^k$ coincide for several index values k, the p^k prefixing such terms may be added, to produce a single term. Consequently, the roulettings provided above are sufficiently unique to define a subsequent bifurcation in the manner of Theorem 4.

If instead of starting as in Theorem 3 with an originally sharp b-plex, we are given an originally undersharp b-plex but with a specific rouletting, then the sharp terms may be bifurcated by a one-dimensional E as above, with a particular rouletting for each such term, these roulettings then rouletted with the originally given coefficients of rouletting, to produce an undersharp result equipped with a specific rouletting. ∎

The Concatenated Thin-Target Thick-Target Model. Here undersharpness through iteration of Theorem 4 is made vivid.

A particle, prepared in state P, is sent through a thick target, conveniently taken as so thick that the particle never gets through: If there is a beam dump, let that be considered as replaced by an opaque succession of further thin layers, subsequently fused. The thick target is supposed to operate as if it were a sandwich of "thin" targets, at each one of which the particle is either stopped, or gets through. If the particle is stopped, some record or signature is left of which thin level stopped it. These signatures are the bins. The projector E for a single thin level is required furthermore to be one-dimensional.

In an actual thick target, there are many signatures possible within one "thin" layer. For example, the layer may be separated in the two dimensions transverse to the beam direction into zones, with specification of the zone wherein the particle has stopped. In each zone, the stoppage may be at any of several chemical species of stopping target particle, subsequently identifiable. But each zone is to be regarded as successively presented to the beam, each chemical population also as successively presented. The idea that presentation of one actual thin layer may be represented by the successive presentation of a large set of simpler thin layers for which stoppage or no is only a 2-test, is a denial of any shadowing effects of one population by another within a thin layer.

The propagation of the particle to thickness t may be represented by replacement of the incident state P by $U(t)P(U(t))^{-1}$, where $U(t)$ is an appropriate unitary operator. However, since $\mathrm{Tr}\ U(t)P(U(t))^{-1}A = \mathrm{Tr}\ P(U(t))^{-1}AU(t)$, it is possible to preserve the format where the incident state is simply P, and the actual acceptors at thickness t are $U^{-1}\cdot U$ transforms of "local" acceptors. It is the already transformed matrices, the ones that work with P directly, which are appropriate to the $\mathrm{Tr}PA$ format: Any unitary propagation is already implicitly included in the foregoing discussion. This is a sort of "Heisenberg picture."

Since the bins have been obtained by successive bifurcation—the conditions of filtration of course being assumed—the theorem concerning bifurcation applies, and the test is undersharp. Calculation of each successive bifurcation provides the rouletting needed for calculation of the next succeeding bifurcation, owing to the assumption that the thin slicing of the target could conceptually be taken down to one-dimensional projections E.

Comment on the Nonuniqueness of Bifurcation. Just as a density matrix P represents all that matters for that aspect of a state needed to determine its response to a test, and not any other details concerning the manner in which the state may have been prepared from an antecedent state, so the b-plex $(A_1,\ldots,A_b)$ represents all that matters for that aspect of a test needed to determine its response to a state, and not anything concerning subsequent tests. In the language of linear time: Density matrices transmit what is important to the future, acceptance matrices work on what is incident from the past. The need to know more than the incident density matrix P, the acceptors $A_1,\ldots,A_b$ of the b-test, and the nature (E,F) of the bifurcation of the bth bin, in order to calculate the $b+1$ new acceptors, illustrates the failure of extrapolation of a (state, test) experiment beyond its test to a further test, without the help of extra information.

The particular rouletting by which the undersharp b-test $(A_1,\ldots,A_b)$ was constructed from sharp b-tests $(A_1^k,\ldots,A_b^k)$, was sufficient extra information.

It is perhaps disconcerting that this is no help in bifurcating an extraneous test, although we could try to force the issue as follows. Any b-plex can be written as a *barycentric* combination of sharp tests, that is, a combination using real coefficients which sum to 1, but which may be negative. It is easy to suggest that bifurcation be defined for extraneous tests by specifying a barycentric decomposition into sharp tests. I have not studied how to manage such a rule so as to keep to nonnegative matrices. If some barycentric extension is, however, taken up, then uniqueness of bifurcation could be lost even for *sharp* tests: Even a sharp test can be written in many ways as a *barycentric* combination of other sharp tests.

10.5 Luxuries. It is important to recognize that the above questions about filtrations and bifurcations are outside the basic domain of discourse here. They need not be resolved. Even a test that destroys the state—meaning one for which bins make sense but the question of what state is left after the kth bin has responded does not—is reasonable and useful in MF.

Recall that MF only seeks matrices P,A so that measured frequencies of activation of the bin to which A must be associated when the test is fed by the procedure that produces the state to which P must be associated, will be fitted by Tr PA, preferably with n small. A procedure defining a state may involve earlier stages which constitute the procedure for producing another state, followed by the laboratory procedure leading to b channels which are otherwise considered b bins of a test, requiring that the *k*th channel be activated, perhaps following this record of activation by further procedures, prior to the application of some test. (States produced without being tested are wasted!) Although the foregoing story mitigates the harsh simplicity of "state meets test," that simplicity is basic. MF is itself already sufficiently defined that in finding all solutions for P's and A's to fit Tr PA's to the measured probabilities, one either gets no solutions, a reasonably unique solution, or an annoying plethora of solutions, without any theory of bifurcation, for example. In the event of a unique solution, we can read off which tests are extraneous, the behavior of anything, even extraneous things, under bifurcation, etc.

It is of course possible to seek to tighten MF, if the general format lead to too many solutions, by guessing that some states are pure, some tests sharp, and that other tests are concocted by an impulsive bifurcative succession of sharp tests, injecting such requirements a priori as constraints on the matrices. Taken together with the undersharpness theorems, such assumptions could render an overall constraint of tests to undersharpness plausible. Or such assumptions could be imposed on some states and tests with others regarded out of bounds for assumptions, with some such tests coming out extraneous.

Another similarly optional luxury would be a requirement that a certain state be the tensor product of two others, or that the density matrices for a certain class of states be of form $P = I \otimes P'$, while those of another class of states be of form $P'' \otimes I$, with yet others of form $P'' \otimes P'$, and still others of general form. Bose and Fermi statistics could also be tried.

The propagation of states in time by Hamiltonians, or with entropy increase, or entropy decrease, are also luxuries: Whether a sequence of states differing only in a coasting time parameter are related through unitary conjugation generated by a Hamiltonian, or otherwise, could be seen by examining the output state matrices without imposing any prior model of motion. Canonical formalism is at best a luxury, perhaps an encumbrance—an expensive luxury, rendering n infinite.

The possible power of the bare Tr PA format without luxuries will hopefully make experimenters eager to gather data for fitting MF without first worrying about theories of the luxuries.

A filling-in of the bare format which would *reveal* some of these conventional structures of ordinary quantum physics approximately, would be more convincing than one obtained by imposing luxuries as constraints of the format. A statistically good small-n fit entirely devoid of any of the usual interrelationships, could be a startling proof of the power of quantum logic, a new sort of tight "factor-analytic" parametrization of some empirical domain, which might yet be puzzling to reduce to new insights.

I should again temper enthusiasm for bare MF by cautioning that there will usually be too many parameters to fit. A priori constraints of course reduce the number of parameters.

11. REPERTORIES

The following results of MB show that the probability data input to MF can suffice to fix a unique solution, of course modulo unitary and antiunitary conjugation, admittedly in a very simple class of cases.

> *Lemma 1: Trace Lemma (MB).* b-plex $(A_1,\ldots,A_b)$ is sharp if and only if it is trace-orthogonal, that is, Tr $A_iA_j=0$ for all $i\neq j$.

A sharp test all of whose acceptors are one-dimensional projections, equivalently a sharp n-plex none of whose matrices is 0, is a *complete test.* This is the one-observable case of Dirac's complete set of commuting observables. A list $(P_1,\ldots,P_n)$ of n states and an n-test $(A_1,\ldots,A_n)$, with Tr $P_iA_j=\delta_{ij}$, will be called a *repertory*.

> *Theorem 5: Repertory Theorem (MB).* The n-plex $(A_1,\ldots,A_b)$ of a repertory is a complete test, and $P_i=A_i$.

The values of Tr P_iA_j are measurable given the procedures for preparing the P_i and for effecting the test $(A_1,\ldots,A_n)$. The fact Tr $P_iA_j=\delta_{ij}$ is learned from empirical probabilities; the n-test need not be known to be sharp or complete in advance. Therefore, repertories are empirically definable, except for guessing n. This n is, however, to be kept minimal for overall fit, corresponding to a refusal to make room for hidden mechanisms in advance of states and tests that require them. For example, enlargement of the matrices of atomic physics upon discovery of electron spin would have been pointless prior to the task of inclusion of spin-dependent effects into the experimental situation (fine structure, sensible exposition of the exclusion principle). When what seemed to be one-dimensional projections must split owing to finer new experiments, that is the time to enlarge the matrices.

We thus fall back on philosophy to bound n from above, but we may have facts to bound n from below: If s states $P_1,\ldots,P_s$ confront an s-test $(A_1,\ldots,A_s)$, and the probabilities Tr P_iA_j are δ_{ij} as for a repertory in $n\times n$ MF, then $n\geqslant s$ (MB).

The notion of repertory associates certain bins of certain tests with states, $A_i=P_i$, even if the test destroys the state, or if the remains of the state are not P_i: What the bin in question "catches" is the same (meaning same matrix) as what some state procedure, possibly unrelated to the test, "produces." If the test is destructive then preparation of state i can nevertheless provide a state like what the ith test bin would have output were the test a filtration: Repertories are technically as useful as "reduction of the wave packet."

Theorem 6: Freezing Modulo Conjugations. The values n and Tr PA for a sufficient set of states P and acceptors A define all the matrices up to an overall unitary or antiunitary conjugation. In particular, the set of all the matrices from all the complete repertories is sufficient: Given the values Tr P_iA_j for all state matrices of all repertories and for all acceptors of all other repertories, and the values of all Tr P_iA for one unknown acceptor and all repertory state matrices, or else given the values of all Tr PA_i for one unknown state matrix P and all repertory test matrices, then all the $P_i = A_i$ are first fixed up to unitary or antiunitary equivalence, and the further information about A or P then fixes A or P in any choice of representation.

Proof. $P_i = A_i = |x_i\rangle\langle x_i|$. Since Tr $P_iA_j = |\langle x_i|x_j\rangle|^2$, the absolute values of all inner products of empirically labeled unit vectors are given: The "empiric label" is the designation by state, as defined by the preparation procedure, or by test and bin. The mathematical content is that there are so many experiments defining complete repertories that *all unit vectors occur* as x_i in $P_i = |x_i\rangle\langle x_i| = A_i$. A well-known theorem of Wigner (1959, Bargmann, 1964) states that a mapping of Hilbert space that preserves the norm must be either a unitary or an antiunitary conjugation. Hence if particular vectors x_i (n-tuples of complex numbers) are found to yield the given $|\langle x_i|x_j\rangle|$ values, then any other vector solution is either a unitary or antiunitary transform of these.

Finally the Tr P_iM or Tr MA_i empiric probabilities for an unknown M give $\langle x|M|x\rangle$ for all x, and thus fix M. ∎

Comment: Freezing with a Finite Number of Experiments. The infinitude of experiments required above is fortunately superfluous. A completely unknown Hermitian matrix M requires only n^2 independent real data to define its elements, so the Tr P_iM or Tr MA_i should fix M, for only n^2 matrices P_i or A_i. Indeed the n^2 one-dimensional projections $P_1 = |x_1\rangle\langle x_1|, \ldots, P_{n^2} = |x_{n^2}\rangle\langle x_{n^2}|$ do define M through Tr P_iM, if the n^2 vectors x_i start with $x_1, \ldots, x_n$, an orthonormal Hilbert-space basis, followed by the $2^{-\frac{1}{2}}(x_i + x_j)$ for $i<j$, then by the $2^{-\frac{1}{2}}(x_i + ix_j)$ for $i<j$.

The one-dimensional projections *from several repertories* are less efficient: They are constrained to give a list where the first n projections are mutually orthogonal as above, but then the next n are also mutually orthogonal, etc. After the n diagonal elements M_{ii} are attained as Tr P_iM from the repertory (taken diagonal for convenience), the trace $\Sigma_i M_{ii}$ is known, whence the n diagonal elements $(U_rMU_r^{-1})_{ii}$ available from each rth repertory for $r>1$ give at most $n-1$ independent new pieces of information. To fix any M, at least n more repertories are needed; indeed $n+n(n-1)$ comes out n^2. This makes at least $n+1$ repertories altogether. By using $U_r = I + \in H_r$ infinitesimally different from I, and H_r an off-diagonal matrix with ones connecting only one particular diagonal spot to all other places, then also other H_r's with i's and $-i$'s placed similarly, I can show that $2n+1$ repertories suffice, but expect that only the trace redundancy really matters, and will take it that $n+1$ repertories suffice.

If M is known to be a one-dimensional projection, then it is determined by the $2n$ real and imaginary parts of the n complex components of vector x, where $M = |x\rangle\langle x|$; normalization and irrelevance of x's phase show that in fact there are only $2n-2$ independent real parameters. Therefore, Tr P_iM or Tr MA_i information for only two sufficiently disparate repertories, should suffice to fix a state or acceptance matrix *already known to be a one-dimensional projection*. Two, not one: one may say a repertory and a *complementary* repertory.

The matrices of a first repertory may be chosen, say, P_i diagonal with 1 in the ith place. The only unitary equivalence yet allowed is the n phases of the basis vectors which render this so. An overall phase is the trivial unitary transformation, affecting no P or A matrix, so only $n-1$ real parameters are undefined. These may be fixed by requiring the *first* matrix P_1' of a second repertory, where Tr $P_i(P_1' = A_1') \neq 0$, all i, to have all its elements positive, in which case the empirical Tr P_iA_1' fix P_1'. Although no parameter is left unfixed for the *eventual* fit by these conventions, the empirical data Tr $P_iA_j' =$ Tr $P_j'A_i$ gained by confronting the states and tests among only two repertories are insufficient to then fix the $P_2', \ldots, P_n'$ completely: A repertory is defined by n^2-n independent real parameters, whereas of the n^2 data Tr P_iA_j', the conditions $P_n = I - \Sigma_{i=1}^{n-1} P_i$ and $A_n' = I - \Sigma_{i=1}^{n-1} A_i$ show at most $(n-1)^2$ of these are independent, leaving at least $n^2-n-(n^2-2n+1) = n-1$ undefined parameters. Whether the data from the mutual experiments with a third repertory suffice to freeze the matrices, I do not know. The "antiunitary" option of course enters when complex elements appear, in that "i" can be replaced everywhere by "$-i$".

In the problem of fixing the second repertory, write $P_j' = |x_j'\rangle\langle x_j'|$, $x_j' = a_{ji}x_i$, to compute the matrix elements $\langle x_i|P_j'|x_k\rangle = a_{jk}^* a_{ji}$, in terms of an incompletely known unitary matrix a_{ji}, with only $|a_{ji}|$ given. Choosing P_1''s matrix elements positive fixes the first a_{j1}. The n phases of a_{j2} must fit orthogonality of the first and second columns, leaving only $n-1$ free phases. Similarly, there are only $n-2$ free phases for the third column,..., leaving only $n(n-1)/2$ phases undefined. Finally, the overall phases of $x'_2,\ldots,x'_n$ do not matter to the $P'_2,\ldots,P'_n$ matrices, $n-1$ inessential phases, leaving $(n-2)(n^2-1)/2$ essential phases. In particular, for $n=2$, this is 0. (This is easier to see from $P'_2=I-P'_1$ being fixed by the convention of positivity for P'_1.) For $n=2$, complex numbers need not appear until a third repertory is considered. But for $n=3$, $\frac{1}{2}$ $(n-2)(n-1)=1$; one essential phase.

If the $\frac{1}{2}(n-2)(n-1)$ phases at the stage of the second repertory are chosen, then each subsequent one-dimensional projection from subsequent repertories will be defined by its empirical relationship with the first two repertories. A wrong choice—and except for "$i\to -i$," phase reversal, there is only one right choice—will eventually lead to an inconsistency, possibly much later, but probably already at the fitting of the third repertory.

But given the right phases, somehow, for two repertories, all other one-dimensional projections are fixed by their empiric relation to the two repertories. The matrices for $n+1$ repertories altogether are then "easily" obtained, if so many repertories can be found, and then any state or acceptor is defined by its empiric relationship to the $n+1$ repertories.

12. VECTOR FORMAT VF, DOT-PRODUCT FORMATS DF_1, DF_2

Regard the $n\times n$ Hermitian matrices as a real n^2-dimensional Hilbert space with Tr $AB\to A\cdot B$ (Section 1.6). Note that $I\cdot I=n$. A basis convenient for some purposes is $I, n-1$ matrices with a single 1 on the diagonal, $\frac{1}{2}(n^2-n)$ with two ones mutually transposed, $\frac{1}{2}(n^2-n)$ with $\pm(-1)^{\frac{1}{2}}$ mutually transposed. A Schmidt process produces orthonormal basis $(e_0,e_1,\ldots,e_{n^2-1})$ with $e_0=n^{-\frac{1}{2}}I$. $e_0e_i=0$, $i\geqslant 1$ signifies tracelessness of those e_i (as matrices). *Except for nonnegativity*, MF is evidently about dot products of vectors. This information supplemented by a determination of which vectors are nonnegative is vector format, VF. I is identifiable by $I\cdot I=n$ and by the central nature of the I axis in the nonnegative cone (*not*, however, a cone of revolution: Appendix C or D). If the vectors are matrices, with nonnegativity determined by determinants, VF is MF in disguise. Theorem 6 indeed shows that designation of the appropriate vectors as nonnegative already limits the notion of equivalence to the usual conjugation. But the dot products are invariant to the larger $O(n^2)$ group; even the $O(n^2-1)$ group of I-fixed rotations is larger (in dimension) than the conjugations if $n>2$ (though they do match for $n=2$). Hence, we have the following theorem.

> *Theorem 7: VF Inexpressible through Dot Products.* Nonnegativity cannot be expressed in terms of dot products or orthogonal invariants (essentially dot products, Weyl, 1946).

Two Dot-Product Formats. Nevertheless this can be circumvented as follows at the cost of enlarging the equivalences to either the $O(n^2)$ or the $O(n^2-1)$ groups. Let any rotate [by either $O(n^2)$ or $O(n^2-1)$] of the nonnegatives be called a *nonnegative shape*. A solution will be a set of vector states, acceptors all lying in *any common* nonnegative shape. This defines dot-product formats $\mathrm{DF}_1, \mathrm{DF}_2$. Theorem 7 no longer obviously applies, and indeed $\mathrm{DF}_1, \mathrm{DF}_2$ can be set forth explicitly. But first expand general matrix M on the e basis, $M_{ab}=\sum_{i=0}^{n^2-1} m_i(e_i)_{ab}$. Work out the formula for $\det(xI+M)$ in terms of the m_i. It will be a polynomial in x with polynomial-in-m coefficients $\gamma_k(m)$, $k=0,\ldots,n$, with $p_n(m)=1$. The nonnegative body is specified explicitly by $p_k(m)\geqslant 0$ conjointly, $k=0,\ldots,n-1$. (Appendix A, Theorem 9): So far this is homework, leading to explicit polynomials in m.

Now, let $x=(x_0,\ldots,x_{n^2-1})$ be any orthonormal list, where $x_0=e_0$ if I is to be fixed, otherwise free. The *nonnegative shape with respect to x* is the set $\{\sum_{i=0}^{n^2-1} m_i x_i : p_0(m)\geqslant 0,\ldots,p_{n-1}(m)\geqslant 0\}$. But $m_i=x_i\cdot M$ is a dot product. Hence $p_k(m_0,\ldots,m_{n^2-1})\geqslant 0$ is in principle explicitly stated in terms of dot products. DF_1 and DF_2 are then "explicitly"

$$x_i\cdot x_j=\delta_{ij}, \qquad x_0\cdot P_a=n^{-\frac{1}{2}}$$

$$\sum_k A_{jk}\cdot x_i=n^{\frac{1}{2}}\delta_{0i}, \qquad P_i\cdot A_{jk}=p_{ijk}$$

$$p_i(x_0\cdot P_a,\ldots,x_{n^2-1}\cdot P_a)\geqslant 0$$

$$p_i(x_0\cdot A_{jk},\ldots,x_{n^2-1}\cdot A_{jk})\geqslant 0$$

Perhaps a nonorthonormal basis e_i', with $e_i'\cdot e_j'=g_{ij}$, could provide easier polynomials p_i', to simplify this orthogonal-invariant approach.

An attempt to define computation entirely in terms of dot products themselves, without the vectors and hence without orthogonal invariance, is the sparse format of MB.

13. THE CASE $N=2$

Separating the Trace Part (n Not Necessarily 2).

Any matrix M is uniquely decomposable according to $M=(1/n)$ (Tr $M)I+M'$, into a multiple of I and its traceless part M'. This corresponds to writing the 0 component separately, $M=\sum_{i=0}^{n^2-1}m_ie_i \to (m_0e_0, \sum_{i=1}^{n^2-1}m_ie_i)$, in VF language. We can split the dot product, $A\cdot B=a_0b_0+(a,b)$. $a=\sum_{i=1}^{n^2-1}a_ie_i$ is the *vector part* of A, a_0 is the *trace part* of A, $a_0=n^{-\frac{1}{2}}\text{Tr } A$.

The compact nonnegative body B is conveniently taken as the set of vector parts of the intersection of the nonnegative cone with the sheet of matrices of trace 1. I can be largely suppressed: Since every state matrix P has trace 1 (0 component $n^{-\frac{1}{2}}$), nonnegativity of a state matrix is rendered by $\rho\in B$, where ρ is the vector part of P. Then Tr PA or $P\cdot A$ becomes $n^{-\frac{1}{2}}a_0+(\rho,a)$. Nonnegativity for acceptor A becomes $a_0\geqslant 0$ and either $a=0$ if $a_0=0$, or $n^{-\frac{1}{2}}a_0^{-1}a\in B$ if not. The requirement that acceptors $A_1,\ldots,A_b$ sum to I becomes that their trace parts sum to the number $n^{\frac{1}{2}}$, $a_{10}+\cdots+a_{b0}=n^{\frac{1}{2}}$, whereas their vector parts sum to the 0 vector, $a_1+\cdots+a_b=0$.

$n=2$. VF is here particularly convenient, because for $n=2$ the compact positive body B *is* $O(3)$-invariant; the compact "shapes" all coincide with B. In fact, the Schmidt-process basis is

$$2^{-\frac{1}{2}}\begin{pmatrix}1&0\\0&1\end{pmatrix},\quad 2^{-\frac{1}{2}}\begin{pmatrix}1&0\\0&-1\end{pmatrix},\quad 2^{-\frac{1}{2}}\begin{pmatrix}0&1\\1&0\end{pmatrix},\quad 2^{-\frac{1}{2}}\begin{pmatrix}0&-i\\i&0\end{pmatrix}$$

by choice of phases; by renumbering these, respectively, e_0, e_3, e_1, e_2, we have $e_i=2^{-\frac{1}{2}}\sigma_i$, in terms of the usual Pauli matrices. The eigenvalues of $\sum_{i=1}^{3}a_i\sigma_i$ are $\pm[(a_1)^2+(a_2)^2+(a_3)^2]^{\frac{1}{2}}$; i.e., $\pm(a,a)^{\frac{1}{2}}$; therefore $A\geqslant 0$ if and only if $a_0\geqslant(a,a)^{\frac{1}{2}}$. The specialization $a_0\to 2^{-\frac{1}{2}}$ produces the trace-1 positive body B, the ball of radius $2^{-\frac{1}{2}}$ in 3-space.

VF for $n=2$ can therefore be described as follows: Each state is represented by a real 3-space vector ρ, with $|\rho|\leqslant 2^{-\frac{1}{2}}$. Each b-test is represented by b nonnegative numbers $a_{10},\ldots,a_{b0}$ with $a_{10}+\cdots+a_{b0}=2^{\frac{1}{2}}$, and by b 3-vectors $a_1,\ldots,a_b$, where $|a_i|\leqslant a_{i0}$, and $a_1+\cdots+a_b=0$. The probability Tr PA that state ρ activate bin k with parameters a_{k0} and a_k is

$$2^{-\frac{1}{2}}a_{k0}+(\rho,a_k)$$

The completely unpolarized state $P=\frac{1}{2}I=2^{-\frac{1}{2}}e_0$ has $\rho=0$. Thus $2^{-\frac{1}{2}}a_{k0}$ is the probability that bin k accept from the completely unpolarized state. So the a_{jk0} can be ascertained prior to seeking vector parts if an empirical procedure is known for preparing the completely unpolarized state, for such tests j as can be coupled with a completely unpolarized procedure.

If $n=2$ is only a hope for study of an unexplored domain, no unpolarized procedure may be established, and the a_{jk0}'s as well as all ρ's and a_{jk}'s would be unknowns in a grand fit to the state–bin probabilities. In the event of 2-tests with a strongly plausible "symmetry" between the two bins, one might wish to try $a_0=b_0$, in which case $a_0=b_0=2^{-\frac{1}{2}}$. Such assumptions of symmetry might run counter to a factual bias, and should be dropped if they fail to produce a solution. For example, if the empirical domain is part of visual perception, symmetry between "up" and "down" is unlikely for a type of subject interested differentially between things in the sky and things on the ground.

Picturing the Case $n=2$. The real 4-tuples (m_0, m) map the 4-space of Hermitian matrices. The nonnegative cone is given by $m_0 \geqslant (m,m)^{\frac{1}{2}}$. If the object is depicted with m_0, time and m, a point in 3-space, then the cone is a movie of events beginning with a dot at the origin and growing outwards with speed 1 as a solid ball, which becomes the ball of states at time $2^{-\frac{1}{2}}$.

The 2-plexes or questions $(A, I-A)$ are indexed by A which lie not only within this growing ball, but also lie in a similar shrinking ball, starting at radius $2^{\frac{1}{2}}$ at time 0. Thus, the set of all such A starts as a point at time 0, grows uniformly as a ball until its radius becomes $2^{-\frac{1}{2}}$ at time $2^{-\frac{1}{2}}$, then shrinks uniformly back to a point at time $2^{\frac{1}{2}}$. This body is the *spindle* of MB. Appendix D partly describes the spindle for general n.

The ball at the central time $2^{-\frac{1}{2}}$ is the ball of states; its 2-sphere surface, the pure states, one-dimensional projections in MF, its center, the unpolarized state $\frac{1}{2}I$.

The 2-plexes are ordered pairs of events with midpoint at the center $\frac{1}{2}I$ of the spindle. Extreme 2-plexes, the sharp questions, are either ordering of diametrically opposite points on the 2-sphere *surface* of the ball at central time $2^{-\frac{1}{2}}$, or of the single pair 0 and I, the spatial center at time 0 and $2^{\frac{1}{2}}$. So a sharp 0-free 2-test has both trace parts $2^{-\frac{1}{2}}$ and vector parts of length $2^{-\frac{1}{2}}$.

14. THE PROGRAM

The problem here is how to search for matrices that satisfy MF for given empirical probabilities p_{ijk}. Ideas that lead to a computer program will be given (Lubkin and Lubkin, 1979b presents a program for $n = 2, 3, 4$), but first the possibility of an analytical solution is discussed in relation to the problem of invariance.

14.1 Invariance. If a problem has 0 or 1 solution, then a computation can arrive at a proof that there is no solution or at the unique solution. If, however, the solutions are more numerous, no *one* may be selected by the proposed computation. MF is therefore ill equipped for computation because for any one solution there is an infinitude of equivalent conjugate solutions. It is therefore desirable to reformulate the problem in terms of conjugation invariants. The device of imposing "irrelevant" further conditions to single out a solution will be regarded as a reformulation of the problem however too devoid of the symmetry which renders things analytically tractable. The part played by a random start in a computer search is, however, an example of getting on with it by in fact using arbitrary further conditions to break through the impasse in a nonanalytic context.

It is for this purpose of fighting invariance by reformulation that formats other than MF were explored. But VF is like MF. DF_1, DF_2 leave the logic invariant under a yet larger group, $O(n^2)$ or $O(n^2 - 1)$, and arrive at a situation in which nonnegativity assumes a complicated aspect, if solution vectors are required. Nevertheless, the DF equations involve only orthogonal invariants, the dot products. A suggestion for computing new dot products from old is sparse format, in MB.

14.2 Search. In the absence of a conjugation-free algorithm, it may be worthwhile to seek a method of searching for a solution to MF by trial and error, suitable for a computing machine, using a random start to break the invariance impasse. Program MATRIXFORMAT for $n = 2$, 3, or 4 and comparison program CLSSCLFORMAT restricted to diagonal matrices (nominally $n \leqslant 99$) are discussed in Lubkin and Lubkin (1979b) and are available from the Computer Physics Communications Program Library. It is probably worthwhile to have trial matrices P_i, A_{jk} chosen identically nonnegative, with Tr $P_i = 1$, and $\Sigma_{k=1}^{b_j} A_{jk} = I$, perhaps by the asymmetrical approach of choosing only $A_{j1}, \ldots, A_{j,b_j-1}$ and putting $A_{j,b_j} = I - \Sigma_{k<b_j} A_{jk}$. Verification of nonnegativity could be done by the n determinantal conditions of Appendix A. But in our programs, nonnegativity was guaranteed by using matrices that were squares of other matrices.

A somewhat arbitrary measure of demerit would be $\phi = \Sigma(\text{Tr } P_i A_{jk} - p_{ijk})^2$, the sum running over those i, j, k for which p_{ijk} is empirically known. A "heteroscedastic" approach is a possible refinement, weights w_{ij} or w_{ijk} being included as factors of the terms in ϕ so as to make statistically better-defined p_{ijk} values have greater importance, or in order to otherwise skew the importance of the p_{ijk} data. Our programs use ϕ with a w_{ijk} option.

The task is to minimize ϕ by trying different lists of P and A matrices. This is to be done for small n first, starting at $n=2$ unless a repertory with larger n is already evident from the data. If the best fit for some n is judged bad, then n is increased by 1. In our programs n is set by the user. ϕ is an inhomogeneous quartic in the real and imaginary parts of the matrix elements of the P's and A's. Those of these real variables not fixed by the Tr $P=1$ and $\Sigma A=I$ conditions are the independent variables x_i. The $\partial\phi/\partial x^i$ could be computed analytically, being inhomogeneous cubics. Of course, $-\partial\phi/\partial x^i$ is the direction of most rapid decrease of ϕ; trying new P,A via $x^i \to x^i - \varepsilon\partial\phi/\partial x^i$, for ε positive, is probably better than searching randomly. For such variation, the Taylor expansion to the quadratic term estimates the variation of ϕ as being

$$-\varepsilon\sum_i \frac{\partial\phi}{\partial x_i}\frac{\partial\phi}{\partial x_i}+\frac{\varepsilon^2}{2}\sum_{ij}\frac{\partial\phi}{\partial x_i}\frac{\partial^2\phi}{\partial x_i\partial x_j}\frac{\partial\phi}{\partial x_j}$$

more briefly, $-\varepsilon\phi_i\phi_i+\frac{1}{2}\varepsilon^2\phi_i\phi_{ij}\phi_j$. This is minimum or maximum for $\varepsilon=\phi_i\phi_i/\phi_j\phi_{jk}\phi_k$. If the quadratic approximation is good, as is likely if a local minimum is close, this should be a minimum, and give an easy estimate of how far to go in the direction of most rapid decrease. If this ε is negative owing to negativity of $\phi_j\phi_{jk}\phi_k$, then the situation is more like getting away from a maximum, and a notion of how large a positive ε should be used in stepping is not provided by the quadratic estimate.

The matrices of the last trial will all be nonnegative. If positive, then a sufficiently small step will not change that. Nonnegativity need not therefore be incorporated into the first decision about where to step to; but after this is tentatively decided, the nonnegativity would have to be rechecked. If violated, a shorter step can be tried. Our matrix-squaring technique provided more definitely for nonnegativity.

It seems a priori unlikely that any of the actual states and tests be pure or sharp, or even that any of the P,A matrices lie precisely on the boundary of the positive cone. If the search procedure outlined tries to head for such a realistic, all-P-and-A-internal solution, it should not seek to cross the determinant-zero boundary of the positives. On the other hand, a spuriously good fit violating nonnegativity may attract a program no better equipped, and some matrices might show at least an early tendency to try to cross the boundary. To cure such an evil, it might be necessary to freeze such matrices, then to later let them loose again and hope they go strongly positive, owing to a better settling of the other matrices.

Especially if the idea of an $n \times n$ matrix fit is all wrong for some empirical domain, then the hope that a search in heading for a reasonable solution will in the main try to avoid the boundary *cannot* be relied on to keep the calculation within the boundary. Also spurious over-the-boundary local minima may attract more strongly than a distant good, all-nonnegative minimum. In that case, the matrices may hit the boundary, and even more extreme surfaces of the boundary in a program which allows motion along the boundary. A tendency for some matrices to have zero eigenvalues may thus mark an insufficiently complete search, rather than excellence of the experimenters in preparing nearly pure states and nearly sharp tests.

Again, specters of badly controlled nonnegativity caused us to bypass these issues by squaring techniques, in our programs.

In such computerized search, the problem of conjugation equivalence may seem not to arise at all: The program would not bother to conjugate its matrices because ϕ would remain unchanged; the stepping would be in quite a different direction, designed to efficiently decrease ϕ. To understand where there is nevertheless a battle against conjugation equivalence, consider the suggestion that search start at all matrices proportional to I and, say, bin-independent within each test: $P_i = n^{-1}I$, $A_{jk} = b_j^{-1}I$, instead of random matrices. An I-proportional point in search space is conjugation-invariant. The directions away from it are not, except for those effecting mere reweighting of the A_{jk} away from equal weights for all bins in one test, so after optimal reweighting, $-\partial\phi$ must fail to point towards a minimum, by being zero. A sufficiently asymmetric choice must be made either by fiat or else by means of (pseudo-) random numbers (which could of course be injected inefficiently through rounding error!). An I-proportional case, even if stationary, is of course not usually the solution. Similarly, the first time an imaginary part becomes necessary, its sign would have to be injected by ex machina noise or by a programmed rule.

14.3 Extreme Formats. Before a fit, it is not known whether empirical procedures in fact provide pure states or sharp tests, or what n to use. Yet near-extremity of procedures given n may be proven by a fit, the fit furthermore rendering the value of n plausible. The P, A matrices of these unmixed or slightly mixed procedures will be close to the boundary and even to subfaces of the boundary of the positive cone. Unless there is such near pinning to the boundary, the solution will display some extra orthogonal-transformation rattling within the positive cone beyond unitary and antiunitary conjugation. An optimal approach to the data would use the interrelationships between the near-extreme "good" experimental procedures and also the other, more mixed procedures, to firm up the whole solution.

It is nevertheless possible to impose *extreme formats* where either state procedures are by fiat pure, or test procedures are by fiat sharp, or both. An actually badly mixed procedure would then have to be represented by a best approximating extreme point, far away from the actually optimal nonextreme point outlawed by the format. In being far off anyhow, it will not greatly enlarge the demerit ϕ for the point to vary from its not-too-good best place. The representatives of badly mixed experimental procedures would therefore be shown up, within an extreme format, by being ill defined, e.g., by having large errors quoted. The procedures themselves could then easily become regarded as unsatisfactory empirically—and their mixity in some sense does render them less informative than purer procedures—and in consequence dropped from the empirical domain and from the program. The loss of weakly determined elements should not radically modify the loci of the better-determined points, but these may even be expected to become much better defined, owing to omission of the errors involved in overlooking mixity of the elements now cast out.

So if there are enough nearly extreme procedures actually in the empirical domain, the nature of the quantum logic could emerge, even under the handicap of an extreme format.

14.4 Apology to Statisticians. Queues. The level of the mathematical statistics involved in asserting a program of minimization of an ad hoc demerit ϕ is undeniably primitive. Also the incorporation of standard statistical notions and methods to evaluate the reliability of a best solution, to quote errors, is obviously wanted.

It may be worthwhile to draw attention to one particular statistical worry. The trials are supposed to be statistically independent. It is nevertheless tempting to define many trials from a single extended queue procedure of form preparation 1, test 1, preparation 2, test 2,...,preparation q, test q, *a queue of length* q. Preparation r specifies how to use what is left as a state, after effecting test $r-1$. The queue may be regarded as split into preparation of a state and execution of a test in q different ways, the split being immediately before test r in the rth way. The history prior to the split, i.e., preparation 1, test 1 and its outcome,...,preparation r, is an elaborate state procedure; the history after the split is an elaborate test procedure, whose bins are specifications of all the outcomes of the subsequent simple tests $r, r+1, \ldots, q$. It may seem wrong to use a single passage through a queue to provide q "independent" data for the whole fit, one datum for each of the q ways of splitting the queue.

Queues were used multiply in our visual perception experiment (Section 15.2), and not to relate the elaborate states and tests referring to the whole histories before and after each split, which would have been too numerous, but to the simple states and tests. In order that the history prior to the immediate preparation before a split have no systematic influence on the succeeding test, that history was randomized.

14.5 Analogy to Factor Analysis. In spite of its need for improvement from old-fashioned statistics, MF or VF may already be a new statistical method. A solution invests some empirical preparations with state matrices or vectors P, and some empirical tests with b-plexes of matrices or vectors A; these mathematical objects entail notions of identity, orthogonality, more generally of angles between procedures, etc. The angles are known only a posteriori, after many trials and subsequent calculation, not through mere description of the empirical methods for effecting the procedures. We get a geometrical depiction of procedures here through experiment and calculation, and that is what people are after who use factor analysis.

14.6 The Statistical Danger of Early Choice of a Representation. It is usual in quantum mechanics to introduce a choice of representation. For example, algebraic relationships assumed to hold between the matrices may lead to a proof that, up to conjugation, a particular representation may be used, with many of the matrices known prior to any empirical trials. The choice of representation is usually dominated by theory. This is reasonable where physical theory antecedent to the quantum approach is respectable. A choice of representation may also be indicated early by the early discovery, through trials, of a repertory, guessing n. The program suggested, on the other hand, is meant to deal with a situation where there is no clear prior theoretical bias to respect, and where no clear repertory presses for an easier sort of analysis.

There may be a virtue of conceptual clarity, nevertheless, in attempting to fix a representation by declaring a certain matrix M to be diagonal, perhaps another one real, etc. But such a procedure may be statistically crude. Indeed, it may become clear only later, from the trials, that M's relationship to the other matrices X, the dot products $M \cdot X$, are only poorly known. By defining standards relative to such a poorly known quantity, the accuracy latent in better-known $X \cdot X'$ dot products could be suppressed. What quantities are most accurately fixed by the data may be hard to anticipate, and may when known turn out to be awkward conceptually.

I have presented a method not tied to prior choice of representation in order to be statistically fair to the data.

15. BACK TO QUANTUM PSYCHOLOGY

There follow thoughts on quantum psychology best presented after accepting the subject as legitimate, and after laying the foundation of model-independent quantum mechanics typified by MF.

15.1 Dreams. A Heisenberg-microscope discussion in quantum psychology provides new thoughts about the Freudian unconscious, and dreams.

The idea is to try to beat the uncertainty principle, by determining which of several slits has been traversed by a particle, yet without spoiling the diffraction pattern. If this can be done, the matrix mechanics behind the uncertainty principle would be shown invalid.

It is supposed that, say, application of MF to a domain of psychological experimentation has generated an association of state matrices to specified procedures of preparation, so that one may speak without vagueness about "states of mind." Of course there are also "bin acceptors of mind"!

Suppose that a particular state of mind is a known linear superposition $y=\sum_i a_i x_i$ of other states, x_i. The x_i are known to be orthonormal, perhaps because they are part of a repertory. These states $y, x_1, \ldots, x_n$ are possible states of mind of a single subject, produced by known procedures of preparation. Let there also be a "diffractive" consequence of the state y sensitive to the phases of the coefficients a_i, not just to the $|a_i|^2$. Then a concurrent demonstration of the value of an observable which chooses between the different x_i must be impossible.

Suppose that the subject can, by introspection, determine which one of $x_1, \ldots, x_n$ describes his state of mind in regard to an x issue, and can perhaps also convey this knowledge to a piece of paper. Then the uncertainty principle could be violated in the diffractive experiment, if after that experiment the subject could faithfully report what his state of mind was in regard to the x issue before application of the diffractive test, by recalling the outcome of his introspective determination. If the earlier introspective outcome were recorded on paper, the recall could be effected by simply reading the paper.

The information on the paper provides the mutually orthogonal attributes not involved in the diffractive test, needed to apply the interference theorem. The introspective record would also do the same thing.

But it does not follow absolutely that quantum psychology therefore does not make sense, because it could be the supposition of self-knowledge of one's attitude to the x issue that is at fault.

The procedure for effecting the x test may well be no more than to ask the subject to mentally select between alternatives $x_1, \ldots, x_n$, a thing so easy that the subject is capable of performing the x-test introspectively. If another procedure has been applied to cast the subject into state y, then that other procedure must nevertheless be incompatible with the subject's ability to perform the x test himself with sufficient force as either to fix a reliable memory of the result, or else with sufficient leisure to allow the production of a paper record.

The closest this can be pushed short of a contradiction is to indeed have an x determination but to fail to remember the result. This already sounds like the Freudian unconscious, or like a beginning of a theory of dreams.

It is puzzling that mental processes of great significance should be unavailable to introspection (Groddeck, 1961). An easy solution to this puzzle is that precisely those mental processes unavailable to "consciousness" most easily go awry through lack of control, hence gain a pathological importance; but I will ignore this easy way out to point out a possible quantum-mechanical solution: The operation of uncertainty principles not only permits models whose functioning necessarily entails forgetting, but even suggests that the "strongest" mental processes are those *most* involved with forgetting and unconsciousness. This follows from the associations of weakness of interaction to classical ontology and of strength to quantum ontology, and from the argument for strength at the center of control. The quantum nature of the center of control should go with lots of uncertainty or unknowability principles, incompatible with the keeping of a record in memory of its operations.

It is easy to lose coherence in an interference experiment by allowing the different x_i to propagate characterizing attributes to things not being tested. Fast techniques may therefore be essential for quantum psychology. A lack of leisure to think about the x issue may be crucial.

15.2 Our Psychological Experiment. These comments about dreams and the unconscious reinforce our enthusiasm for doing psychological experiments suitable for study by MF. The easiest case is $n=2$, a "polarization." This makes two-way flip–flop psychological effects attractive: an $n=2$ empirical domain at least has no sharp b-test with $b>2$. I will continue to explain our motivation for such an experiment, much as I proposed at the May 1973 London, Ontario conference (Lubkin, 1976), then comment on the actual experiment and on our negative result.

Analogy to a plane-polarizing filter for light suggests that we seek a dichotomic illusion $(A(\theta), B(\theta))$ with a continuous cyclic parameter θ such that $A(\theta+\text{const})=B(\theta)$ and $B(\theta+\text{const})=A(\theta)$:

$$- +\qquad -^{+}\qquad \pm\qquad +_{-}\qquad + -$$

The symmetry implied for the two views perhaps makes the word "illusion" inappropriate. In any case, the Necker cube (Attneave, 1971; Gregory, 1968, 1966) of Figure 2, which can be seen as a cube in perspective in two ways, is such a display. It has the property that if the right-hand one of the two corners internal to the hexagonal boundary of

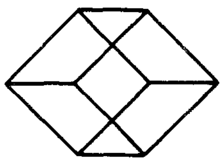

Fig. 2. The Necker cube at $\theta=0$.

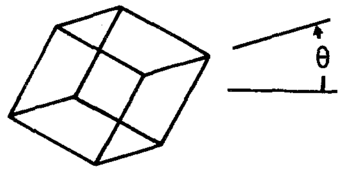

Fig. 3. The Necker cube at another θ.

the drawing is seen "high" and then the actual planar drawing is turned within the plane, and if the high internal corner is tracked continuously, one passes to the state of mind in which the original concrete figure, now restored by a planar congruence, is seen *the other way*. Thus "const" $=\pi$; the planar turning of a sheet of polaroid gets to the other polarization with "const" $=\pi/2$ instead, in the optical analogy.

Our experiment used eight equally spaced θ-rotated views of the Necker figure $\pi/8$ apart: call these *cubes*. The cubes were projected on an HP-1317 cathode ray tube display screen under the control of a MODCOMP minicomputer. The points of one cube were once and for all precomputed on the UWM Univac installation and prepared in scrambled order, and seven rotates of this one cube were also Univac prepared, so that the minicomputer had only to make a data transfer from its own fast memory in order to transmit a new view to the screen. Thus the whole view seemed to appear at once—and because of the scrambling, the points in the various edges were in fact not filled in in any neat ordering. I note this because readers familiar with minicomputer displays may imagine a figure slowly "painted" in—whatever faults our experiment may have had, this was not one of them!

A run went as follows. The first thing visible on the screen was a small, centered circle, and a somewhat more intense blot or "point" of light. The subject S was asked to run the point into the fixed circle, by manipulating a joystick. It is the time *that* took which selected which of the eight cubes appeared as the first display, assuring a random choice of first view: poor human control over time made S our analog and principal randomizing device, both in this minor matter of assuring a random first view and in the major matter of assuring independently random subsequent views.

With the appearance of a cube, the point disappeared, S now finding the small circle to be under control of the joystick. S's next task was to conceive the cube three-dimensionally, to decide which of the two cube vertices closest in the actual planar figure to the center ("central corners") appeared to S to be three-dimensionally closer, and to guide the moving circle to this apparently closer corner. Collision of the circle with one of the two central corners constituted the objective means of reportage by which S's manner of seeing the cube three-dimensionally was transmitted to fast storage. Also transmitted at this time was the choice of cube itself and the time taken to make that choice, rounded off to tenths of a second —but we have not had occasion to seriously use this timing information, except that if S dawdled too long (30 sec) the run was terminated. In this case the collection of a next report effected termination, with the tardy report being dropped.

This circle-corner collision also brought up the next view: a new random cube would appear with a newly centered circle not controlled by the joystick, the last joystick position, now eccentric, reverting to the intense-point marker. From here on, the story repeats: S would have to move the point back to the central circle in order to get control of the circle, so as to be able to report S's three-dimensional interpretation of the new figure. In this possibly clumsy way, S would not be misinterpreted *automatically* as reporting the same answer, when the same cube view happened to reappear.

If a run was not terminated first by S taking too long to report a view, it was terminated at 1024 views. It is only at termination of a run that data were transferred from minicomputer fast storage to tape.

A typical subject would take 1 sec, or a little less, per view.

When a new view came up of the same cube, this of course was evident by reappearance of the one-point marker and retreat of the small circle to the center, but there was also apparent a fleeting (and thoroughly satisfactory) "blink," probably corresponding mostly to the reloading of the 4096-point HP-1317 memory, a roughly 0.01-sec procedure.

We were not interested in making distinctions between subjects, not even basically in studying visual perception: we only wished to learn whether a psychological experiment would reveal large quantum coherence effects by being much better fitted by MF than by CF. Hence we had no reason to concern ourselves with choice of subjects and no notion of separation of subjects into a distinguished class and a class of control subjects. I think this aspect of our work seems unbelievably wrong to some literal-minded biologically trained people.

Our main result is negative: we found no evidence for the superiority of a quantum-mechanical model. This makes analysis of fine details uninteresting, so my report is about making the negative situation simply evident (as in Lubkin and Lubkin, 1979a and 1979b), then commenting on weaknesses of our experiment.

It is the seeing of a cube *and* the reporting of a view which we took to define a state. Thus we had 16 states. The *next* display of a cube was taken as a test, hence we had 8 2-tests, the view reported being regarded the outcome, making each bin number b indeed 2. The number of independent probabilities is then 128. Except for the first and last views in a run, each view thus served both as state and test, an experimental design already denoted previously as a queue. Perhaps this is the only *sensible* way to index our experiment in the (state, text, outcome) manner, in which case it may be discourteous to readers to so tediously set it forth, but it is not the *only* way to do it. For example, consecutive *pairs* of views could have been taken as the elements of a half-as-long queue, to give us 256 states and 64 4-tests, and a need for much more data than we have to claim a reasonable determination of probabilities. In some reductions, we weighted trials (i.e., state-test transitions) inversely as the time recorded in the test report, using a heteroscedasticity feature of our MATRIXFORMAT, CLSSCLFORMAT programs (Lubkin and Lubkin, 1979b), but with no interesting consequence.

The quantum-logic dimension (Section 9.3, Theorem 2) for $n=2$ is $d(2)=77$, but $d(3)=192$, $d(4)=353$. Hence except for $n=2$, the 128 independent probabilities are fewer than the number of effective fitting parameters. But our count of 128 independent probabilities could be inflated, if the division of the continuum into 8 states is already too fine, implying needless interpolation. We are up against the difficulty that MF's numerous parameters may too easily fit anything. This is an elementary counting argument known to us in advance, to show that our experiment was marginal.

The total number of trials in our whole data base was 30746. An $n=2$ approximate quantum fit achieved an r.m.s. deviation of theoretical from experimental probabilities of 0.080, whereas a corresponding $n=4$ fit with diagonal matrices achieved a similar r.m.s. value of 0.082. Both these r.m.s. values represent fits to the data so close that some of the noise is being fitted. But their similarity shows that the quantum fit is no better than that classical fit which shares about the same number of essential fitting parameters. This shows that we have no sign of interesting quantum effects, so we have not tried to fish through the noise for them.

Why did we work so hard on a marginal experiment? For two reasons: First, even very gross effects can escape attention until they are deliberately sought: witness the story of the large violations of parity conservation in weak decays. Second, we felt that the concreteness of a real though marginal experiment would draw more attention to the issue of possible quantum correlations in everyday life than the mere theoretical proposal, even if nothing amazing was discovered by the experiment.

Our experiment has other weaknesses.

If S sees the cube in a definite way, that probably means that he has enquired of some partly coherent state which way it is, and has so applied a test, even though the experimenter has not requested him to do so. It is desired that the subject drift "mechanically" during displays and be "tested" only when explicitly requested to record a mode of seeing the cube. If the technique is slow, it will be difficult for *S* to avoid the introduction of *S*'s own ad hoc tests. This indicates that the displays should be somehow too fast to "think about," yet slow enough, of course, to see. I am writing here as if *S* were at least two collaborating mental entities, an elemental seer too simple to fail to have interesting phase relationships (hopefully!), and a thinker, the agent of the memory, which applies reality-defining tests to the output of the seer. The thinker and the external experimenter are both applying tests to the seer, and the experimenter is trying to apply his tests fast enough so that their effect will not be swamped by the noise engendered by the tests introduced by the ordinary thinker. There is the extra difficulty that the experimenter's tests are transmitted to the seer by the very thinker the experimenter is trying to circumvent, but the success of ordinary hypnosis indicates that verbal external direction of the thinker is not impossible. In brief, it is desirable to pick up *S*'s response rapidly and without unduly bothering *S* with busy work. We have *S* playing with a joystick, doing complicated things with moving points and circles. Perhaps the experiment could be improved by visual-perception experts, equipped to record eyeball positions.

Professor Richard Warren of our Psychology Department anticipated another weakness. He expected that *S* would usually prefer one of the two interpretations for quite ordinary reasons. Thus, if one interpretation would have the top face exposed while the other would have the bottom face exposed, the first would be preferred because we more often look down upon cubes than up. This is of course a weakness from the standpoint of our program, because it has each test coming out in a way independent of the preceding state, and all probabilities close to 0 and 1, unlike illustrations where an acceptor significantly fails to commute with a state matrix. And indeed a subject did tend to respond uniformly to each given figure, though not entirely so. I could usually, while watching behind *S*, anticipate the response, and I fear became somewhat annoyed with *S*'s for being so simply predictable!

If one looks at a cube for a while, its three-dimensional interpretation will usually reverse spontaneously, a generally well-known phenomenon. This fact seems to deny the too-uniform response trouble we have had, but our measurement addresses mainly the *first* response, so there is no contradiction. The near balance between the two interpretations seemingly implied by reversal had made us hopeful that we would be studying something sensitive; if so our technique has not developed that sensitivity.

15.3 Ladder to Consciousness? The experimental discovery of some mental quantum interference would be directly interesting. It would begin a quantum-amplitude structure behind everyday probabilities, connecting quantum ontology more directly to common experience. Experiment with an internal observer yet with interference might be feasible in empirical domains where an observer is *not* complex.

A mathematical description of psychological mechanics behind ontology could teach us new things about reality, perhaps leading to alternatives to space–time: The problem of consistency of perceptions of Berkeley and Borges should be regarded as open to study within the prereal framework of quantum theory. Darwinian evolution partly settles this: it is useful to the organism to apprehend perceptions only insofar as they are in fact consistent; yet this language is somewhat circular, "usefulness" presupposing purpose and time whereas the problem is to *generate* purpose and time (and so space). But Darwinian argument has also suggested association between consciousness, nonswitching mechanism, and quantum logic, which may be wrong, so it is not altogether empty.

If this association between consciousness as prime control and quantum interference is right, then a possible ladder to consciousness exists. One starts anywhere in psychology with a program to seek quantum interference—I have suggested visual perception as an example. The experiments then reveal some details of the state space. Other phenomena are allowed to leak in, and the state space description grows. One learns where the interference effects are strongest. This points to the center of control, and hopefully displays its mode of operation.

The ladder to consciousness and a ladder to the mechanism for generation of ontology may even be one and the same.

16. UNIVERSALITY; QUANTUM MATHEMATICS

How interrelated are the various disciplines? Especially, how immune is mathematics from extramathematical invention? Quantum logic at once reveals tighter *interrelationships* yet formalizes new *lack* of relationships.

16.1 Lack. In order not to worry about the smallness of $\hbar$ making quantum effects small, it helps to lose the notion that all things live in one great Hilbert-space universe, and that any experiment belongs to a technical test over this one Hilbert universe. Bohr's complementarity, especially in contrast to the idea of superselection rule, reveals a plurality of incompatible Hilbert spaces in ordinary quantum mechanics *within* physics.

Thermodynamic variables and the classical jelly-raster construction provide further precedents. For me, the least compelling argument is that the burden is on the other side, that an empirical situation need not be considered so specially until the details show that to be so.

So maybe to quantize a non-standard-physics field with equanimity, you first argue that the field is *entirely* cut off from standard quantum physics! So much for lack.

16.2 Interrelationships. You can apply quantum logic to the data analysis of *any* subject where the facts factor into states and tests; all domains have at least this in common.

If the observer is overdelicate with the observed system so that only the system's own physical state seems to matter, the tests only discovering facts about the states without disturbing them, you will not find against classical ontology. Destructive tests that cut deeper may be overlooked because their significance may be incomprehensible in a classically ontological context. Thus the requirement that a state be easy to copy led us from economics to psychology. A study of purpose, goal, utility, morality, more formally of control, would be central to both psychology and economics: A deep approach to control could fuse these disciplines.

Economics could be many-body psychology, but there could also be distinct quantum mechanisms (Section 6.2) cutting across both. Some quantizations might be slavishly discovered by using MF on real experiments, others on the basis of thought experiments or guesses that lead to successful predictions. I am indeed jumping the gun somewhat with the claim that Freudian notions provide a bias for quantum psychology and with the argument from development of *depth* of control through Darwinian evolution.

There are goals in *every* field of study, not just economics and psychology. In this sense, any abstract science of goals is primitive to every discipline. Perhaps there is an as yet nonexistent science of goals to be extracted by means of quantum logic from empirical psychology and economics, goals being commonly understood as the essence of these subjects.

16.3 Mathematics. Of all fields, the one most claimed aloof is mathematics. Even if most new mathematics arises from applications, the axiomatic distillates do seem aloof. But this aloofness is lost when we seek to weigh the relative importance of different axiomatic systems. The axioms are only an interim wall between the vague but important goals and the sharp conclusions that sometimes flow from the axioms.

What would it be like without the wall? That would be saying that you cannot talk sensibly about relative importance of anything, not even of mathematical axiomatic systems, except in reference to a consciousness or control which selects for importance. There is competition, a budget of sorts, so the issue is at once mathematical and economical! Maybe this can be done all internal to mathematics. It would involve different incompatible mathematical realities. This phrasing cheapens quantum ontology to suggest the ordinary situation in mathematics, of plural axiomatic topics. Isolating the invariant reality of a topic from its provincial axiomatic form and interrelating distinctly axiomatized topics is, say, the "theory of categories".

The metamathematical commonplace, that any transmitted mathematical treatment is a finite row of symbols, may be compared to the sequential digital storage of a computer memory, perhaps also of a live memory, to draw a parallel between mathematics and psychology. Contrast with the picture of infinite sets finitely discussed, the "Skolem paradox." Mathematics is the psychological output of the world, drawn in most distinct form. The distinctness is that of symbols. It is shared by the bin disjunction of quantum logic, where such distinctness coexists with unknowability principles.

Perhaps the mind is already so well represented in mathematics and the study of mathematics is already so well developed in category theory and metamathematics that there is already there a structure of quantum psychology, hidden because the state-test experience of a Darwinian mind is not apprehended there. If so, then the smart way to build a machine that thinks may be through understanding where to splice state-test empiricism with category theory. The stupid way is to cheat by looking at real psychological experiments to see where quantum correlations appear before they are understood.

How can one study category theory as if it were a living mind, by confronting states with tests to get concrete statistical data, for input into a quantum-logic reduction? One cannot. How can a student produce a lab report without doing an experiment? By cheating: There is a way! If we can calculate what probabilities we ought to get, we can then cheat and invent spurious data. Identifying states tests and probabilities, relevant somehow to purpose, may be enough of a purpose analysis of a theoretical realm even without made-up random data. The impossibility of random output of a mathematically controlled analysis, even one about probability randomness and quantum ontology, may be as irrelevant to a mathematical study of purpose as the finitary quality of language is irrelevant—hopefully—to the mathematics of infinity.

16.4 Rotation. "Looking at an issue from a different angle" might have a literal quantum-logical interpretation. Inventive intelligence could be a basis-rotating capacity. Fuzziness of the old-basis thought is compensated by storing results in memory. This is separating a theorem from its proof.

Marx brothers' humor: the activation of basis rotation with paradoxically little reward. The other Marxist, Talmudic, lawyers' dialectics: inducing basis rotations by trying to oscillate imperfectly between "yes" and "no."

16.5 Recapitulation: a Thinking Machine. By "thinking machine" I mean a device that thinks like a person, not just a device that answers questions. If the nucleus of human thought is a quantum control unit, then a stored-program computer is not such a device.

Indeed, you have not built a quantum-thinking machine unless it gives different answers to the same question: A consistent computer does not think. The inconsistency is not to be just put in by a meaninglessly randomizing device; noise in itself is bad and even noise generated by quantum operation is not in itself useful. The noise that accompanies quantum operation, however, diagnoses that depth of probing is distorting the thing probed. Depth is good, distortion bad; the best *compromise* therefore allows distortion for the sake of depth. The distortion is ontologically shattering, so it shows up in a reality as randomness.

The argument that the stored-program computer is not a quantum thinker is most convincing when the computer is totally without randomization. If, ignoring this, the argument is to be refuted by using randomization, that must reach to *answers*: Monte Carlo programming to get essentially definite answers will not do. Quantum strength of thinking reaches to unpredictability of the answers. Since the answers reach us, the ontological branching of a thinking machine does involve our own reality.

A thinking machine should not be too silly about giving different answers to the same question. It should be sure that $1+1=2,\ldots$*except*: It should have a chance to decide, on its own initiative, to study the purposes revealed by the definitions that lead up to $1+1=2$, and then by questioning the purposes it should perhaps say that, if the field has characteristic 2 then $1+1=0$, *also* that then only 0 and 1 need be defined making "2" superfluous, hence making $1+1=2$ ungrammatical, hence intolerable. Since any subject goes back to purposes the deep answers to even apparently straightforward questions may strike back to shifting sands. One also wishes to operate at a shallow level, to see the definite consequences of questionable assumptions: a stored-program competence is also useful.

So reliable switches are *only auxiliary tools*, the reliability being good, though diagnostic of a lack of depth of interaction. Switching, reflex units in neural organization are useful *peripheral apparatus.*

Depth of interaction rather than a classically ontological world image should have been primary in Darwinian evolution. All living forms possess sophisticated control mechanisms, even unicellular organisms, and most of the functional chemistry of multicellular organisms is organized within single cells. Just as the digestion by vacuoles gullet and enzymes is developed in the single cell and only slightly adapted in multicellular forms, so nervous organization is likely an extension of the apparatus of control in unicellular forms.

Such organization is necessarily intracellular for unicellular forms. Once such intracellular control is primitively established, it will be elaborated and modified, but the control itself is not likely to be replaced by a new mode of control; mutations make revolutions only through large losses. So it is more plausible that major neural function resides within cells than that cells are simple switching elements.

This has been reworked here because, since it is hard to understand the quantum-logical aspect of a switching system, it is important, if black-box quantum psychology is not to be ruled out a priori by brain structure, to see *how to question* the idea that the brain is a switching system.

As soon as one gets out of a cell, one is likely in the peripheral area, and the control center does best if its peripheral devices *are* reliable switches. So exchanges *between* cells should, like a telephone system, use good switches. Examining the pulses exchanged might not be getting closer to studying brain function than examining telephone conversations. The core of control may escape. Quantum logic may tell more about control, even if the wall between state and test is moved outside the body, than nerve studies without quantum logic. Whole-organism psychological experiments somehow seem first, even if eventually quantum logic does get applied to a cell or ganglion as black box.

The motive for addressing the cell comes from doubt about where a new quantum ontology can fit. Yet the ordinary world is not classically causal! We fabricate machines to transmit our purpose faithfully, and work at this; it is not automatic. The engineering science that has grown up around these machines strives towards the faithfully reproducible. The culminating Laplacian world view of absolute classical ontology is a Freudian rationalization of this line of endeavor.

In fact, quantum wave-packet spreading is largely with us. Deciding where to go on the basis of a real Stern–Gerlach experiment demonstrates this. Thus, "up" sends us through one door, "down" sends us through another, and the whole story sends the mean coordinate of our wave function for our center of mass through or into some wall—a classical limit of little value. This already broad shattering of ontology inherited from atomic physics is so much with us that there is little Laplacian anticipation of a definite story of events in the future. A control mechanism may exist in the *vacuum of definition* together with its quantum ontology; there is no existing classical definiteness to be lost.

The definiteness of events in sequence is, of course, a definiteness of memory, of the past. The definiteness of memory is usually also confirmable in the "objective world" because nonconfirmable memories are of no Darwinian adaptive value.

Can we build something with its own quantum control center to live in its own branching ontology, in a way that the consequences of this branching will somehow not escape us? But we generate new branching in atomic physics experiments, and we know what that does to us: We branch in correlation, but see that only in the stochastic look that a series of such experiments leaves in our memory (Lubkin, 1979). So the only "strange" aspect, but the diagnostic one, of a thinking machine will be that its behavior will not be entirely predictable. Even this will not seem odd because in this it will be like a cat! A live cat.

The diagnosis of quantum control in such a machine could be by a quantum fit of the probabilities through, say, MF. For diagnosis of quantum control in a black box, it is necessary for it to show nontrivial probabilities in its responses to multiple-choice b-bin "questions," which also subsequently reveal underlying quantum coherence upon analysis.

ACKNOWLEDGMENTS

David Finkelstein's encouragement is deeply appreciated. Some of his parallel work should, however, also be mentioned: His independent interest in quantum psychology is prior to mine, he has emphasized the incidental character of probabilities in quantum mechanics (Finkelstein, 1963) by working with the statistical sharpness of a long run (a viewpoint I, however, avoid; Section 6.2 here and Lubkin, 1979), and he has worked on the incapacity of quaternionic quantum theory to deal with compound systems (Section 9.2 and Lubkin, 1979). Paradoxically, Finkelstein disapproves of plural reality.

Professors C. A. Hooker and W. L. Harper of the University of Western Ontario encouraged the proposal of the Necker-cube experiment (Lubkin, 1976) at the Statistics Conference held there May 1973, by running open sessions. Professor R. A. Northouse is thanked for our use of his Robotics and Artificial Intelligence Laboratory at UWM in doing the experiment. Our key program MATRIXFORMAT (Lubkin and Lubkin, 1979b) was significantly improved because of a brief criticism by Kenneth Baclawski.

The enthusiasm of then Milwaukee friends Steven Fulling, George Diderrich, Robert Shanny, and James Wyly, and Marilyn Levine, and the patience of economists at the workshop seminar of Figure 1, are acknowledged.

Most significant was the collaboration of my wife Thelma in formative discussions, in organizing the workshop seminar, finding mathematical errors, and later in programming.
Joyce Miezin is thanked for her artistic typing of the 1973 manuscript and this update.

APPENDIX A: POSITIVITY

As texts on matrices usually treat positivity only for real symmetric matrices, the reader may find the elementary treatment here useful. *Positivity* in the strong sense, all eigenvalues >0, is treated first, then nonnegativity separately.

Theorem 8: Positivity. The following conditions are equivalent, for an $n\times n$ Hermitian[6] matrix M:

(1) The eigenvalues are all >0.

(2) $x\neq 0$ implies $\langle x|M|x\rangle > 0$.

(3) The $k\times k$ Hermitian matrix obtained by erasing the last $n-k$ rows and columns of M has positive determinant, for each $k=0,\ldots,n-1$.

(4) The n determinants of the matrices obtained by successively striking out one row and corresponding column of M one at a time in some *arbitrarily* chosen order, are all positive.

(5) All determinants of square matrices obtained by deleting rows and corresponding columns of M are positive.

(6) $\exists$ nonsingular S such that $M=S^{\dagger}S$.

(6′) $\exists$ positive Hermitian $M^{\frac{1}{2}}$ such that $M=(M^{\frac{1}{2}})^2$.

(7) $\exists$ basis $(x_1,\ldots,x_n)$ such that the n^2 matrix elements M_{ij} of M are given by $M_{ij}=\langle x_i|x_j\rangle$.

Proof. Condition (2) implies (1): For each eigenvalue λ, choose x a normalized eigenvector, $Mx=\lambda x$ and $\langle x|x\rangle=1$. Then $\langle x|M|x\rangle=\lambda$, and from 2, $\langle x|M|x\rangle>0$. Hence $\lambda>0$.

Condition (1) implies (2): Because M is Hermitian, there is an orthonormal basis $(x_1,\ldots,x_n)$ of eigenvectors of $M, Mx_i=\lambda_i x_i$. If $x=\Sigma_i a_i x_i$, then $\langle x|M|x\rangle=\Sigma_i|a_i|^2\lambda_i>0$.

Condition (1) implies (6), (6′): Let $Sx_i=\lambda_i^{\frac{1}{2}}x_i$, where $\lambda_i^{\frac{1}{2}}$ is the positive square root, and extend S linearly, i.e., if $x=\Sigma_i a_i x_i$, then $Sx=\Sigma_i a_i\lambda_i^{\frac{1}{2}}x_i$. Then $S^{\dagger}=S$, and $M=S^2=S^{\dagger}S$.

Condition (6) or (6′) implies (2): I.e., S nonsingular and $x\neq 0$ imply that $\langle x|S^{\dagger}S|x\rangle>0$. Indeed $\langle x|S^{\dagger}S|x\rangle=\langle Sx|Sx\rangle>0$, since Sx is a nonzero vector.

[6]Condition (2) for *arbitrary* square M itself implies that M is Hermitian, but a matrix with n nonorthogonal eigenvectors fulfilling condition (1) is not Hermitian.

Hence (1), (2), (6), (6′) are equivalent; call (1), (2), or (6) *positivity*. It remains to prove that (3), (4), (5), and (7) are equivalent to positivity.

Positivity implies (5): $\det M$ is the product of M's eigenvalues, hence >0; this is the null deletion case of (5). To prove positivity of the determinant of M' obtained from M by deleting the ith rows and columns of M for all $i \in$ Set, observe that the $\langle x|M|x\rangle > 0$ computed using nonzero x whose components x_i for $i \in$ Set vanish, are one–one with equal values $\langle x'|M'|x'\rangle$ where $'$ signifies deletion of vector and matrix components indexed in Set. But the x' so obtained are general nonzero vectors in the vector space after deletion of components, whence $\langle x'|M'|x'\rangle > 0$ satisfies form (2) of positivity for M'. Hence M' is a positive matrix in the Set-deleted space. Hence $\det M' > 0$.

Condition (5) obviously implies (3) and (4).

Condition (4) reduces to (3) by renumbering.

Condition (3) implies (1): It is given that

$$m_{11} > 0, \qquad \begin{vmatrix} m_{11} & m_{12} \\ m_{21} & m_{22} \end{vmatrix} > 0, \qquad \begin{vmatrix} m_{11} & m_{12} & m_{13} \\ m_{21} & m_{22} & m_{23} \\ m_{31} & m_{32} & m_{33} \end{vmatrix} > 0$$

etc.; including $\det M > 0$. By adding a multiple (by factor $-m_{21}/m_{11}$) of the first row of M to the second row of M, a new matrix is obtained with element 0 in place 21. Such a "row operation" changes none of the n given determinants. Now add the complex conjugate amount (factor $-m_{12}/m_{11}$) of the first column to the second column, to restore hermiticity and make place 12 zero. If M is considered a product IMI with unit matrices on the left and right, the row operation may instead be regarded as being performed on the left factor I, the column operation on the right factor I. If the left I is changed thereby to A, then the right I is changed to $A^\dagger$. Thus, the process may be written $M \to AMA^\dagger$. That $(AMA^\dagger)^\dagger = AMA^\dagger$ follows from the identity $(XY)^\dagger = Y^\dagger X^\dagger$; the notation $AMA^\dagger$ shows concisely that the row and column operations together preserve hermiticity.

Next, write $A = I + B$, and consider $M \to (I+\xi B)M(I+\xi B)^\dagger$, where ξ varies from 0 to 1. The effect is still a row and a column operation leaving the n determinants unchanged, but now M varies *continuously* to $AMA^\dagger$, passing through Hermitian matrices in such a way that the eigenvalues, real because $(I+\xi B)M(I+\xi B)^\dagger$ is Hermitian, vary continuously.

Similarly annul the elements in the 31, 41,...,n1 places (and in the 13,...,1n places).

By performing row operations in which a continuous multiple of *row 2* is added to later rows together with the conjugate column operations, all elements in places $k2$ (and $2k$) are also rendered 0 for $k > 2$. This process can be continued to produce a diagonal matrix. Each newly isolated diagonal element is proved positive by observing that the next positive determinant is the product of the previously isolated positive diagonal elements and the newly isolated diagonal element. Therefore the process concludes by producing a positive diagonal matrix.

To see that the *original* matrix M is positive, observe that since the real eigenvalues of M had to vary continuously through real values (the eigenvalues of the varying matrix) to final positive values, any originally negative value would have had to cross through 0. This would have rendered the $n \times n$ determinant zero; instead the original nonzero value of this determinant does not vary. QED: (3) implies positivity.

A corollary of this portion is that, if the n nested determinants are given only *nonzero*, then the sign *changes* starting from a plus sign, proceeding to the 1×1 determinant, then the 2×2 determinant, etc., are in one–one correspondence with the *signs* of the elements of the final diagonal matrix. The number of such reversals is the number of negative eigenvalues of the original matrix.

Positivity implies (7): If the original basis is denoted $(e_1, \ldots, e_n)$, then $M_{ij} = \langle e_i | M | e_j \rangle$. From version (6′) of positivity, $M_{ij} = \langle e_i | M^{\frac{1}{2}} M^{\frac{1}{2}} | e_j \rangle = \langle M^{\frac{1}{2}} e_i | M^{\frac{1}{2}} e_j \rangle$. Thus the $x_i = M^{\frac{1}{2}} e_i$ fulfill (7). They are linearly independent because the e_i are, and $M^{\frac{1}{2}}$ is nonsingular.

Condition (7) implies positivity: Given a basis $(x_1, \ldots, x_n)$, to find that the $\langle x_i | x_j \rangle$ constitute a positive matrix, let S be the nonsingular linear transformation such that $x_i = Se_i$, where $(e_1, \ldots, e_n)$ is the original orthonormal basis. Then $\langle x_i | x_j \rangle = \langle Se_i | Se_j \rangle = \langle e_i | S^\dagger S | e_j \rangle$. These are the matrix elements of positive matrix $S^\dagger S$. ■

Nonnegativity. Condition (7) is weakened to nonnegativity by replacing "basis" by "list of n vectors."

The nonnegative combination $\Sigma p_i M_i$ of nonnegative Hermitian matrices $M_i, p_i \geqslant 0$ is nonnegative (e.g., via Part 2 of Theorem 8); the nonnegative Hermitian matrices form a convex cone of one nappe dropped from the 0 matrix. The positive Hermitian matrices of course form the interior.

It does *not* follow from $M \geqslant 0$ and $M \neq 0$ that $M > 0$, e.g., $\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$. Nevertheless we say $A > B$ when $A - B > 0$; $A \geqslant B$ when $A - B \geqslant 0$. $>$ and $\geqslant$ are strong, weak real vector-space orderings, even though $\geqslant$ is not "$>$ or $=$".

A matrix may have all eigenvalues $+$, or 0, or $-$, or $+0$, or $+-$, or -0, or $+-0$, already seven specific cases; "nonnegative" is less specific and lumps cases $+$, 0, and $+0$.

The n nested determinantal conditions for positivity might suggest that any n exactly similar weak inequalities be equivalent to nonnegativity, but no: $\begin{pmatrix} 0 & 0 \\ 0 & -1 \end{pmatrix}$ has nested determinants 0,0 working down from the upper left. One can get correct conditions by applying positivity to $M+\varepsilon I$ for $\varepsilon>0$, $\varepsilon\to 0$. But the following, stimulated by a conversation with E. P. Wigner, is easier.

Let $\det(xI-M)=p(x)=\sum_{k=0}^{n}p_k x^k$ be the characteristic polynomial of a Hermitian matrix M.

> *Theorem 9: Nonnegativity of* $-M$. $M\leqslant 0$ is equivalent to $p_k\geqslant 0$, $k=0,\ldots,n$.

Proof. Given $M\leqslant 0$, to show $p_k\geqslant 0$. $p(x)$ is the product of factors $x-\lambda$, where λ is an eigenvalue of M, hence $\lambda\leqslant 0$. Thus, $x-\lambda=x+|\lambda|$; since only nonnegative numbers appear in these linear factors, the p_k obtained from multiplying them out are indeed nonnegative.

To show the converse, assume $p_k\geqslant 0$, and note that $p_n=1$, always. Then if $x>0$, $\sum_{k=0}^{n}p_k x^k>0$. Hence there is no positive root. ■

The coefficient p_k can be seen to be the sum of all k-down diagonal minors of matrix $-M$, as follows: In expanding $\det(xI-M)$, x occurs precisely k times, when in the several products of n matrix elements, precisely k diagonal spots are selected to give the x factor, hence the rest of each such product is selected from the matrix $-M$ with the k rows and columns through those selected spots deleted precisely in the manner of a minor determinant. Since $M\leqslant 0$ is equivalent to $-M\geqslant 0$, we can refer directly to the nonnegative matrix $-M$, to get the following theorem.

> *Theorem 10. Sums of Minors.* A Hermitian matrix is nonnegative if and only if the n sums of its k-down diagonal minors, $k=0,\ldots,$ $n-1$, are nonnegative.
>
> More specifically, the product of the $x+|\lambda|$ for $M\leqslant 0$ has ν factors x if ν is the multiplicity of 0, and a product of $n-\nu$ factors $x+|\lambda|$ with $|\lambda|>0$, which multiplies out to a polynomial of degree $n-\nu$ with all coefficients positive. Hence, the terms of low degree of $p(x)$ are zero, but then terms of degree ν or more are positive: $p_k>0$ for $k\geqslant\nu$, $p_k=0$ for $k<\nu$. In words, there are no gaps once the positive coefficients start; in the minors language, for a nonnegative matrix, the low-order sums of diagonal minors are positive, the high-order ones are zero, with no intermixing of zeros among the positives.

> *Corollary.* A real polynomial p with nonnegative coefficients with gaps must have nonreal complex roots.

Proof. Otherwise its roots would all be $\leqslant 0$. The diagonal matrix of these roots repeated according to their multiplicity would be a Hermitian $M \leqslant 0$ whose characteristic polynomial would be p, contradicting the no-gap result above. ■

APPENDIX B: EXTREMITY OF ONE-DIMENSIONAL PROJECTIONS AMONG THE STATES

Lemma 2. The space of two-sided nonnegative variation (MB) of one-dimensional projection P is spanned by P itself.

Proof. Suppose $P+\lambda X$ nonnegative for real λ with $|\lambda|$ sufficiently small. Choose a basis so that $P_{11}=1$, other $P_{jk}=0$. The diagonal elements $(P+\lambda X)_{kk}$ with $k \geqslant 2$ are simply λX_{kk}. Since diagonal elements of a nonnegative matrix are nonnegative, $\lambda X_{kk} \geqslant 0$ whether λ is positive or negative, hence $X_{kk}=0$, $k \geqslant 2$. To see that also $X_{jk}=0$ for $j \neq k$, note the nonnegativity of the 2×2 minor determinant, $(P+\lambda X)_{jj}(P+\lambda X)_{kk}-(P+\lambda X)_{jk}(P+\lambda X)_{kj} \geqslant 0$. This boils down to $-\lambda^2 |X_{jk}|^2 \geqslant 0$, hence such $X_{jk}=0$. This leaves X_{11} as the only allowed two-sided variation, proportional indeed to P. ■

Theorem 11: Extreme States. The extreme points in the compact convex body of nonnegative matrices of trace 1 are the one-dimensional projections.

Proof That the one-Dimensional Projections Are Extreme. Even the one-dimensional variation of Lemma 2 is forbidden by fixity of the trace. ■

Proof, Less Computational, of Theorem 11. A nonnegative matrix of trace 1 not a one-dimensional projection can be diagonalized to display it as a mixture of one-dimensional projections. Were some one-dimensional projection $|x\rangle\langle x|$ not extreme, then invariance of the body B of nonnegative trace-1 matrices to unitary transformation $A \to UAU^{-1}$ would show that *any* one-dimensional projection $|Ux\rangle\langle Ux|$, U arbitrary, is also not extreme, leaving *no* extreme point. But the nonempty compact convex body B is the convex completion of its set of extreme points, so there are some. ■

Proof. , Less Computational, of Lemma 2. Let d be the dimension of the space $X(P)$ of two-sided variation of one-dimensional projection P.

The restriction that the variation not alter the trace cuts the dimension to $d-1$. Extremity of P in the trace-1 positives, Theorem 11 shows this dimension, however, to be 0. Thus $d=1$. Since one dimension is already provided by $X=\lambda P$, no other variation is possible. ■

APPENDIX C: UNITARY INVARIANCE AND THE NONNEGATIVE CONE

It was argued in Section 11 that a sufficient number of probability data will freeze an $n\times n$ matrix solution up to unitary or antiunitary conjugation, $M\to UMU^{-1}$ or $(UMU^{-1})^*$, i.e., that if a solution containing enough independently placed repertories (at most $2n+1$, probably only $n+1$) exists, all other solutions are conjugate, in the sense given. The number of real parameters effective in changing the matrices without changing the physics of state-bin probabilities, is dim $SU(n)=n^2-1$. Aside from dot-product conditions in DF_2 which admit the larger group of effective action $O(n^2-1)$ of dimension $\frac{1}{2}(n^2-1)(n^2-2)$, the only conditions on the matrices are nonnegativity. It follows that the subgroup of $O(n^2-1)$ that preserves the nonnegative cone consists precisely of those particular orthogonal transformations given by $U\cdot U^{-1}$ or $(U\cdot U^{-1})^*$ conjugation.

This can be seen more directly. Let us speak briefly of the $O(n^2-1)$ actions as "orthogonal transformations," the unitary or antiunitary transformations, as "conjugations."

To show that any orthogonal transformation that leaves the nonnegative cone invariant is a conjugation: The general orthogonal transformation is one-to-one. In being real-linear, it preserves convex combinations. Therefore the extreme points of a convex body are mapped one-to-one onto those of the orthogonally transformed image. If that image is to coincide with the original body, the mapping permutes the extreme points.

Although the condition that nonnegative matrices map into nonnegative matrices allows the image to be a subset of the original whole cone, it is easy to see that this cannot in fact be a *proper* subset: The $O(n^2-1)$ in question acts in the sheet of trace-1 matrices; the nonnegative body in question may be taken as B, the set of nonnegative matrices of trace 1, compact, and of finite (n^2-1) volume. Since orthogonals preserve volume, the image subset must be a convex subset of no less volume, and so must be all of B.

The extreme points of B are the one-dimensional projections $|x\rangle\langle x|$, x running over the unit vectors in complex n-dimensional Hilbert space, the $|x\rangle\langle x|$ in one-to-one correspondence with the complex rays in Hilbert space. Furthermore, $|x\rangle\langle x|\cdot|x'\rangle\langle x'| = \mathrm{Tr}\ |x\rangle\langle x|x'\rangle\langle x'| = |\langle x|x'\rangle|^2$ is, being a dot product, preserved by the orthogonal transformation. Therefore, the transformation induces a transformation of unit vectors in Hilbert space which preserves $|\langle x|x'\rangle|$. That such a transformation is necessarily of form $|x\rangle \rightarrow U|x\rangle$ or $|x\rangle \rightarrow (U|x\rangle)^*$, with a common unitary U for all x in Hilbert space, is the same well-known theorem of Wigner that introduced the conjugations in Section 11.

When $x \rightarrow Ux$ or $(Ux)^*$, $|x\rangle\langle x| \rightarrow U|x\rangle\langle x|U^\dagger = U|x\rangle\langle x|U^{-1}$ or $(U|x\rangle\langle x|U^{-1})^*$. There are enough one-dimensional projections $|x\rangle\langle x|$ to form a vector-space basis for the real n^2-dimensional space of Hermitian matrices, on which the orthogonals act real-linearly. Therefore, the extension of $U \cdot U^{-1}$ or $(U \cdot U^{-1})^*$ from the one-dimensional projections is unique. Since $M \rightarrow UMU^{-1}$ or $(UMU^{-1})^*$ is such an extension, it is the only one.

It is nonnegativity of the physical probabilities, the invariance of Tr $PA = P \cdot A$ owing to these dot products being the physical data, and the *real* linearity of MF, which limit the transformation theory to conjugations if P, A run over sufficiently many one-dimensional projections, according to the present proof; not the further evident requirement that the probabilities add to 1 or complex linearity. Of course the norm structure referred to as the base of all this is that of complex Hilbert space.

APPENDIX D: THE SPINDLE, $N \geqslant 2$

The spindle, the set of $A \geqslant 0$ with $I - A$ also $\geqslant 0$, depicts the 2-plexes. A lurid description was given in Section 13 for $n = 2$. The situation for general n is made more vivid here by describing the planar sections of the spindle, through the O–I axis.

Lemma 3. Acuteness. If A, B are positive, then $A \cdot B > 0$.

Proof. Replace A, B by $A' = UAU^{-1}$ and $B' = UBU^{-1}$, U unitary, so that $A' = \mathrm{diag}\ (a_1, \ldots, a_n)$ is diagonal. $A \cdot B = A' \cdot B'$ and $A' \cdot B' = \sum_i a_i (B')_{ii}$ is a sum of products of positive numbers. ∎

Corollary. If two nonnegative vectors A, B are trace-orthogonal, i.e., $A \cdot B = 0$, then they *both* lie in the boundary of the positive cone, i.e., det $A = 0$, det $B = 0$.

Proof. Draw the plane through O, A, and B. The nonnegative vectors in this plane form an intersection of two convex bodies, the plane and the nonnegative cone, and is therefore itself convex. It is also an intersection of two cones, hence a cone. Hence, it is a convex cone in the plane OAB.

Such a figure is obviously a sector between two bounding rays. The angle between the two bounding rays is at most a right angle; otherwise slightly interior rays at more than a right angle would violate Lemma 3. Hence angle AOB is less than a right angle, unless both A and B are on the bounding rays. ■

Axial Planar Intersections. The spindle may be described as the intersection of the nonnegative cone with the figure obtained by adding the constant vector $I = n^{\frac{1}{2}}e_0$ pointwise to the nonpositive cone (the second nappe of the cone of two nappes which extends the nonnegative cone). Adding $I = n^{\frac{1}{2}}e_0$ is 0-translation through distance $n^{\frac{1}{2}}$.

Pass a plane through O and I. The plane intersects the space orthogonal to e_0 in a line on which lie two unit vectors, m and $-m$. The plane is the set $\{xe_0 + ym\colon x,y \text{ real}\}$, and $e_0 \cdot m = 0$; i.e., matrix m is traceless. (x,y) are Cartesian coordinates in this plane. $0 = (0,0)$, $I = (n^{\frac{1}{2}},0)$ are extreme points, so the segment $\{(x,0)\colon 0 \leqslant x \leqslant n^{\frac{1}{2}}\}$ lies in the spindle, the other $(x,0)$ do not. This segment lies within the sector between two extreme rays of nonnegatives, the positive cone within the plane, of opening angle at most a right angle. When the second nappe extending this sector is 0-translated by amount $n^{\frac{1}{2}}$, and the intersection is taken, a parallelogram is produced: Figure 4.

Let the angle from the 0-axis to the bounding positive-cone ray nearest m be $\theta(m)$. It is determined from the eigenvalues of m: Indeed, the point $(1,y)$ on the ray corresponds to the largest y for which $e_0 + ym$ is positive. If the eigenvalues of m are $\lambda > \cdots > -\mu$, in decreasing order, then those of $e_0 + ym$ are $n^{-\frac{1}{2}} + y\lambda > \cdots > n^{-\frac{1}{2}} - y\mu = 0$. Hence, $y = \tan\theta(m) = n^{-\frac{1}{2}}\mu^{-1}$. Similarly, $\tan\theta(-m) = n^{-\frac{1}{2}}\lambda^{-1}$. From $m \cdot m = 1$, the sum of squares of eigenvalues is 1; then Tr $m = 0$ shows that there are both

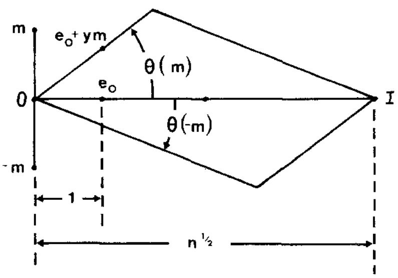

Fig. 4. Intersection of the spindle with a plane through the OI line, a parallelogram with $\theta(m) + \theta(-m) < \pi/2$.

positive and negative eigenvalues.

$$\tan[\theta(m)+\theta(-m)]=\frac{\tan\theta(m)+\tan\theta(-m)}{1-\tan\theta(m)\tan\theta(-m)}$$

$$=\frac{n^{-\frac{1}{2}}\mu^{-1}+n^{-\frac{1}{2}}\lambda^{-1}}{1-n^{-1}\mu^{-1}\lambda^{-1}}$$

$$=n^{\frac{1}{2}}\frac{\lambda+\mu}{n\lambda\mu-1}$$

The extreme points of the *parallelogram* are its four vertices. Of these, O and I are extreme points of the spindle. Internal points of the parallelogram are of course also internal to the spindle. In our planar section only the other two vertices, where the positive and translated-negative cone boundaries cross, might be extreme in the larger spindle.

If one of these points is A, the other is $I-A$. Extremity of A in the spindle is equivalent to sharpness of the question $(A,I-A)$, equivalent to $A\cdot(I-A)=0$ by the trace lemma (MB and Lemma 1). Therefore, whether or not the vertices of the parallelogram other than O,I are extreme *in the spindle* depends on whether $\theta(m)+\theta(-m)$ is a right angle or an acute angle. A right angle corresponds to infinite tangent, $n\lambda\mu=1$, an acute angle to $n\lambda\mu>1$. A is extreme if and only if the parallelogram is a rectangle.

Because there are no extraneous 2-tests, an extreme A must be simply a projection of dimension d. Tr $A=d$, so $A_0=n^{-\frac{1}{2}}d$ is the x coordinate of the vertex on the $\theta(m)$-inclined ray. Since $\tan\theta(m)=n^{-\frac{1}{2}}\mu^{-1}$, the vertex is $(n^{-\frac{1}{2}}d, n^{-1}d\mu^{-1})$. The point $I-A$ is $\left(n^{-\frac{1}{2}}(n-d),\ -n^{-1}(n-d)\lambda^{-1}\right)$. Orthogonality $A\cdot(I-A)=0$ yields only $n\lambda\mu=1$ again, but the condition that these sum to $I=(n^{\frac{1}{2}},0)$ gives $d\mu^{-1}=(n-d)\lambda^{-1}$, or $\lambda\mu^{-1}=(n-d)/d$. Multiplying by $\lambda\mu=n^{-1}$ yields $\lambda=[(n-d)/dn]^{\frac{1}{2}}$. The matrix m is obtained from the projection A by removing the trace part and normalizing; thus there are only two distinct eigenvalues, $[(n-d)/dn]^{\frac{1}{2}}$, d-fold, and $-[d/(n-d)n]^{\frac{1}{2}}$, $(n-d)$-fold.

In other words, the intersection of the spindle with a plane through the three points O, I, and another extreme point A, is a rectangle which when scaled up a factor $n^{\frac{1}{2}}$ can be drawn as follows: Draw a circle of diameter n, and draw a diameter marked O at the left and I at the right. Erect a one-sided perpendicular to the diameter at distance d from the left. The intersection of this perpendicular with the circle marks "A," a third corner of the rectangle. The various cases $d=0,1,2,3,4,5$, for $n=5$, are shown superposed on one circle in Figure 5, although for any particular

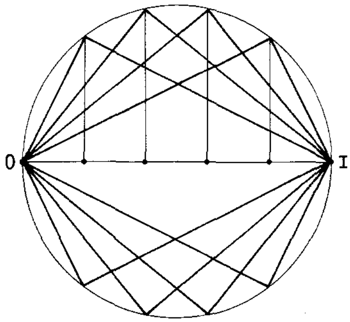

Fig. 5. Shapes of the planar sections of the spindle illustrated for $n=5$ and $(d,n-d)=(1,4)$, $(2,3)$, $(3,2)$, $(4,1)$. Cases $(0,5)$ and $(5,0)$ degenerate to the segment OI.

plane through O, I, and an extreme point A (and so also the extreme point $I-A$), only one of these rectangles applies.

Were these rectangles $O(n^2-1)$-rotated about the O–I axis freely, and the convex body taken, call it the *twirl*, one would get for its planar intersection the convex polygon determined by *all* the corners (including O, I) in each plane, a decagon in the illustration. The twirl is the smallest convex *axisymmetric* figure that contains the spindle properly. The *disparity* between the twirl and the spindle itself illustrates the lack of $O(n^2-1)$-invariance of the spindle, $n \geqslant 3$. To mark this irrelevance, the six nonrectangle edges of the decagon are *not* drawn in Figure 5.

The one-dimensional projections, the extreme rays of the nonnegative cone, have $d=1$, so lie at the rather large angle $\theta = \cos^{-1} n^{-\frac{1}{2}}$ from the O–I axis, while never being more than a right angle apart from each other.

APPENDIX E: DIMENSIONS OF VARIOUS FIGURES

Figure	Real Dimension	Complex Dimension
$n\times n$ complex matrices	$2n^2$	n^2
$n\times n$ Hermitian matrices	n^2	
$n\times n$ nonnegative Hermitian matrices, the nonnegative cone	n^2	
$n\times n$ unitary matrices $U(n)$	n^2	
$SU(n)$, $SU(n)$ mod its center $\cong U(n)/U(1)$	n^2-1	
all possible acceptors	n^2	

Figure	Real Dimension	Complex Dimension
states, or "density matrices," or the compact positive body B at trace 1, or space of cone elements or rays of the positive cone	n^2-1	
pure states, or Hilbert-space rays, or one-dimensional projections, or extreme points of nonnegative cone, or homogeneous space with congruences $U(n)$ and isotropy $U(1)\oplus U(n-1)$	$2n-2$	$n-1$
Hilbert-space vectors	$2n$	n
sharp questions or 2-plexes (E,F) of dimensions $(d,n-d)$, or homogeneous space with congruences $U(n)$ and isotropy $U(d)\oplus U(n-d)$, or d-dimensional extreme points of spindle	$n^2-d^2-(n-d)^2 = 2d(n-d)$	
general 2-plexes, or questions; the spindle	n^2	
sharp b-plexes $(A_1,\ldots,A_b)$, dim $A_i=d_i$, $\Sigma d_i=n$, or congruences $U(n)$ with isotropy $\oplus_i U(d_i)$	$n^2-\Sigma d_i^2=\sum_{i\neq j} d_i d_j = 2\sum_{i<j} d_i d_j$	
repertories ($d_i=1$ above)	n^2-n	
all b-plexes	$n^2(b-1)$	
undersharp b-plexes	$n^2(b-1)$	
$O(N)$	$\frac{1}{2}N(N-1)$	
$O(n^2)$	$\frac{1}{2}n^2(n^2-1)$	
$O(n^2-1)$	$\frac{1}{2}(n^2-1)(n^2-2)$	
boundary of the positive cone	n^2-1	
boundary of the compact nonnegative body B	n^2-2	
intersection of the boundary of the positive cone with the boundary of the I-translated negative cone	n^2-2	
logics in the $n\times n$ matrix format for s states and t tests with $b_1,\ldots,b_t$ bins	n^2m-s+1 where $m=s-1+\sum_{j=1}^{t} b_j$	
classical logics with $n\times n$ diagonal matrices, s states and t tests as above	$n(m+1)-s$ m as above	

APPENDIX F: DISCONNECTEDNESS OF THE CONJUGATIONS IN RELATION TO DISCONNECTEDNESS OF $O(n^2)$ OR $O(n^2-1)$

$U(n)$ and therefore the $U \cdot U^{-1}$ conjugations, are connected (Chevalley, 1946); this is readily apprehended by changing U to I through successive one-parameter "rotations" preserving unitarity. The difference between two antiunitary conjugations is a unitary conjugation, and it is easily verified that the antiunitary, complex conjugation cannot be effected by a unitary conjugation; thus, the antiunitary conjugations form a distinct connected component. This well-known two-component structure of the conjugations appropriate to quantum mechanics in MF without superselection rules is related to the two-component structure of the orthogonals, only when $n \equiv 2$ or 3 mod 4:

The real-vector-space basis of n^2 Hermitian matrices of Section 12 has $\frac{1}{2}n(n+1)$ real matrices and $\frac{1}{2}n(n-1)$, pure imaginary. Complex conjugation, on this basis, therefore reverses the signs of precisely $\frac{1}{2}n(n-1)$ matrices. The determinant of either the $O(n^2)$ or $O(n^2-1)$ real-linear operations which effect the corresponding transformation in either DF is therefore $(-1)^{n(n-1)/2}$. For $n = 2, 3, 4, 5, 6, 7, \ldots$, this is, respectively, $-1, -1, 1, 1, -1, -1, \ldots$, a sign reversal every second step.

The orthogonal groups also have two components, the orthogonals in the component of the identity element being characterized by determinant 1, the other component, -1. Thus, when $(-1)^{n(n-1)/2} = 1$, the antiunitary conjugations belong to the same connected component of the orthogonals as the unitary conjugations, for the other n, however, to the other component. Restriction of the orthogonals down to the conjugations only, produces two connected components; these both being carved out of the full orthogonal component of identity only when $(-1)^{n(n-1)/2} = 1$.

REFERENCES

Aharonov, Y., and Susskind, L. (1967a). *Physical Review*, **155**, 1428.

Aharonov, Y., and Susskind, L. (1967b). *Physical Review*, **158**, 1237.

Attneave, F. (1971). *Scientific American*, **225**, No. 6, December, p. 63 and cover.

Bargmann, V. (1964). *Journal of Mathematical Physics*, **5**, 862.

Birkhoff, G., and von Neumann, J. (1936). *Annals of Mathematics*, **37**, 823. Reprinted in von Neumann (1962).

Bohr, N. (1934). *Atomic Theory and the Description of Nature*. Cambridge University Press, Cambridge.

Bohr, N. (1950). *Science*, **111**, 51.

Borges, J. L. (1967). *A Personal Anthology*. Grove Press, especially "A New Refutation of Time."

Chevalley, C. C. (1946). *Theory of Lie Groups: I.* Princeton University Press, Princeton, New Jersey.

Colodny, R. G. (ed.). (1972). *Paradigms and Paradoxes*. University of Pittsburgh Press; Pittsburgh, Pennsylvania.

Dirac, P. A. M. (1947). *Quantum Mechanics*, Third Edition. Oxford University Press; Oxford.

Einstein, A. (1949). In *Albert Einstein Philosopher–Scientist*, Vol. II., P. A. Schilpp, ed. Harper & Bros., New York.

Einstein, A. (1953). *The Meaning of Relativity*, Fourth Edition. Princeton University Press, Princeton, New Jersey, especially p. 3.

Epstein, S. T. (1960). *Nuovo Cimento*, **X16**, 362.

Everett, H., III. (1957). *Reviews of Modern Physics*, **29**, 454.

Finkelstein, D. (1963). *Transactions of the New York Academy of Sciences*, **25**, 621.

Finkelstein, D., Jauch, J. M., Schiminovich, S., and Speiser, D. (1962). *Journal of Mathematical Physics*, **3**, 207.

Frazer, W. R. (1966). *Elementary Particles*, Prentice-Hall, Englewood Cliffs, New Jersey, p. 28.

Gardner, M. (1973). *Scientific American*, **229**, No. 1, July, p. 104.

Giles, R. (1970). *Journal of Mathematical Physics*, **11**, 2139.

Gleason, A. M. (1953). *Journal of Mathematics and Mechanics*, **6**, 885.

Gregory, R. L. (1966). *Eye and Brain*. McGraw-Hill, New York.

Gregory, R. L. (1968). *Scientific American*, **219**, No. 5, November, p. 66.

Greidanus, J. H. (1971). *Proceedings Koninklijke Akademie van Wetenschappen Series B*, **74**, 445.

Groddeck, G. (1961). *The Book of the It*, tr. V. M. E. Collins. Knopf and Random House, New York, p. 12.

Hegerfeldt, G. C., Kraus, K., and Wigner, E. P. (1968). *Journal of Mathematical Physics*, **9**, 2029.

Jauch, J. M. (1968). *Foundations of Quantum Mechanics*. Addison-Wesley, Reading, Massachusetts.

Jauch, J. M. (1973). *Are Quanta Real? A Galilean Dialogue*. Indiana University Press, Bloomington, Indiana.

Landau, L. D. (1927). *Zeitschrift für Physik*, **45**, 430.

Latham, R. E. (1951). Lucretius' *The Nature of the Universe*. Penguin Classics, Toronto.

Lubkin, E. (1960). *Nuovo Cimento*, **X16**, 1098.

Lubkin, E. (1964). *Nuovo Cimento*, **X32**, 171.

Lubkin, E. (1968). *Bulletin of the American Physical Society II*, **13**, 590.

Lubkin, E. (1970). *Annals of Physics*, **56**, 69.

Lubkin, E. (1971). *Bulletin of the American Physical Society II*, **16**, 67.

Lubkin, E. (1974a). *Journal of Mathematical Physics*, **15**, 663.

Lubkin, E. (1974b). *Journal of Mathematical Physics*, **15**, 673. MB designates (1974a and b).

Lubkin, E. (1976), in *Foundations of Probability Theory, Statistical Inference, and Statistical Theories of Science, Vol. III*, W. L. Harper and C. A. Hooker, eds., Reidel, Dordrecht.

Lubkin, E. (1977). *Abstracts of Contributed Papers, 8th International Conference on General Relativity and Gravitation, August 7–12, 1977, University of Waterloo, Ontario, Canada*, duplicated for the conference by the Faculty of Mathematics.

Lubkin, E. (1978). *Journal of Mathematical Physics*, **19**, 1028.

Lubkin, E., and Lubkin, T. (1979a). *Bulletin of the American Physical Society II*, **24**, 54.

Lubkin, E., and Lubkin, T. (1979b). *Computer Physics Communications*, **16**, 207.

Lubkin, E. (1979). *International Journal of Theoretical Physics*, **18**, 165.

Mackey, G. W. (1963). *Mathematical Foundations of Quantum Mechanics*. Benjamin, New York.

Mielnick, B. (1980). *Journal of Mathematical Physics*, **21**, 44.

Mirman, R. (1969). *Physical Review*, **186**, 1380.

Mirman, R. (1970). *Physical Review D*,**1**, 3349.

Mirman, R. (1979). *Foundations of Physics*, **9**, 283.

Moravec, H. (1979). *Analog* XCVIX, No. 2, February, p. 59.

Neumann, J. v. (1927). *Göttinger Nachrichten*, 245, reprinted in von Neumann (1961).

Neumann, J. v. (1955). *Mathematical Foundations of Quantum Mechanics*, tr. R. Beyer. Princeton University Press, Princeton, New Jersey.

Neumann, J. v. (1961). *Logic, Theory of Sets and Quantum Mechanics*, Vol. I of *John von Neumann, Collected Works*. Pergamon Press, New York.

Neumann, J. v. (1962). *Continuous Geometry and Other Topics*, Vol. IV of *John von Neumann, Collected Works*. Pergamon Press, New York.

Rolnick, W. (1967). *Physical Review Letters*, **19**, 717.

Satir, P. (1961). *Scientific American*, **204**, No. 2, February, p. 108.

Schrödinger, E. (1935). *Naturwissenschaffen*, **48**, 52.

Schrödinger, E. (1945). *What Is Life?* Cambridge University Press and MacMillan, London.

Schrödinger, E. (1958). *Mind and Matter*. Cambridge University Press, Cambridge.

Stapp, H. P. (1971). *Physical Review D*, **3**, 1303.

Stapp, H. P. (1972). *American Journal of Physics*, **40**, 1098.

Weyl, H. (1928). *Zeitschrift für Physik*, **46**, 1.

Weyl, H. (1946). *The Classical Groups*. Princeton University Press, Princeton, New Jersey.

Weyl, H. (1949). *Philosophy of Mathematics and Natural Science*. Princeton University Press, Princeton, New Jersey.

Wheeler, J. A. (1957). *Reviews of Modern Physics*, **29**, 463.

Wick, G. C., Wightman, A. S., and Wigner, E. P. (1952). *Physical Review*, **88**, 101.

Wick, G. C., Wightman, A. S., and Wigner, E. P. (1970). *Physical Review D*, **1**, 3267.

Wigner, E. P. (1952). *Zeitschrift für Physik*, **133**, 101, footnote 2, p. 102.

Wigner, E. P. (1959). *Group Theory*. Academic Press, New York, pp. 233–236.