Quantum Phenomenology

Allan F. Randall
Dept. of Philosophy, York University
Toronto, Ontario, Canada
research@allanrandall.ca, http://allanrandall.ca

Abstract. Starting with the Descartes' *cogito,* "I think, therefore I am"—and taking an uncompromisingly rational, rigorously phenomenological approach—I attempt to derive the basic principles of recursion theory (the backbone of all mathematics and logic), and from that the principles of feedback control theory (the backbone of all biology), leading to the basic ideas of quantum mechanics (the backbone of all physics). What is derived is not the full quantum theory, but a basic framework—derived from *a priori* principles along with common everyday experience—of how the universe of everyday experience *should* work if it operates according to rational principles. We find, to our surprise, that the resulting system has all the most puzzling features of quantum physics that make physicists scratch their heads. Far from being "bizarre" and "weird", as is usually thought, the strangest paradoxes of quantum theory turn out to be just what one ought to expect of a rational universe. It is the classical, pre-quantum universe of the nineteenth century that has irrational, mystical components. The quantum-mechanics-like theory that is developed is, furthermore, most compatible with the strictest, most uncompromisingly rationalist of the standard interpretations of quantum mechanics, those which add no ad hoc elements to the theory, and which generally trace their history to the relative state formulation of Everett (also called the "many worlds" interpretation). These interpretations take the universe to be quite literally describable as a quantum wavefunction. As with any project this far-reaching in scope, I confess I have had to make some working assumptions along the way. I have attempted to isolate these, and clearly label them as points of possible future revision—they are marked in the text with an asterisk (*).

Introduction

At some point in your life, dear reader, you have probably been smitten with the sense of wonder. Perhaps not just now, as you sit here with your mind crowded by other things (not the least of which is to understand these words), but at some time in the past, you have probably looked around in utter astonishment at the sheer existence of the world. “What *is* this that I am experiencing, anyway? How can it possibly be? What am I, and why is there anything at all, rather than just *nothing*?”

If you have never been tormented by such thoughts, the words that follow will mean little to you. The search for the answers to such questions is called “metaphysics”, and the approach to metaphysics I shall explore here is sometimes called “phenomenology”. Broadly speaking, it was invented by René Descartes,[1] although it was further developed by many others, including Edmund Husserl,[2] who is sometimes called its father. Phenomenology looks at the real question

[1]Rene Descartes. *Discourse on Method and Meditations on First Philosophy* (3rd. Ed.), Donald A. Cress (Trans.). Hackett, Indianapolis, 1637, 1641, 1993.

[2]Edmund Husserl. *Cartesian Meditations: An Introduction to Phenomenology*, Dorion Cairns (Trans.). Kluwer Academic, Dordrecht, The Netherlands, 1929, 1950, 1995.

that lies behind all other questions asked in science and philosophy: what is the explanation for my current conscious experience? Everything else boils down to some variation on this question. Everything else you know—the external world, your past memories—is known to you as perceptions laid before your consciousness. As such, any science, whether it be metaphysics, physics, biology or economics, must be properly grounded as an attempt to explain consciousness.

Any such exploration must start with some variation on Descartes' *Cogito*—"I think, therefore I am"—and build on this in a series of movements that step away from the *Cogito*. We will not require absolute certainty at every step, so long as each movement is motivated by an attempt to understand immediate consciousness, and not some prior prejudice about what that is. This will lead us first to the search for an object *within* consciousness that can be experienced with complete clarity. We will find that mathematical truth provides the basic model for clarity of experience, and that any attempt to understand consciousness will therefore be mathematical (at least until some other model of clarity is found). We will therefore want to understand what mathematics is, from the standpoint of conscious experience.

With mathematics as our primary tool, we will examine our immediate consciousness carefully, and try to build a clear mathematical model that captures the essential structure of experience. This will lead us to rediscover the basic concepts of feedback control theory.[3,4,5,6] Further investigations will allow us to begin to build a proper phenomenological conception of our "world", leading to a rediscovery of the basic principles of our most fundamental physical theory, quantum mechanics.[7,8]

Cogito Ergo Sum

Phenomenology necessitates putting aside all prejudice as to just *what* one's conscious experience is. Initially, we must not assume even the existence of an external world. The world is not known directly to us, only certain conscious perceptions are, so the "world", if there is any such thing at all, must for the time being be put aside, placed in parentheses or "bracketed", as an unjustified prejudice. This is Descartes' method of radical doubt, called by Husserl the *epoché*. It is always possible, Descartes tells me, that my notions of an external world might be completely

---

[3]Aristotle. "Nicomachean Ethics," In: *The Basic Works of Aristotle*, pp. 935-1126. Random House, New York, c. 335-322 BC, 1941.

[4]Norbert Wiener. *Cybernetics: or Control and Communication in the Animal and the Machine* (2nd Ed.). MIT Press, Cambridge, MA, 1948, 1961.

[5]W.R. Ashby. *An Introduction to Cybernetics*. John Wiley & Sons, New York, 1956.

[6]William T. Powers. *Behavior: The Control of Perception*. Aldine Publishing, Hawthorne, NY, 1973.

[7]Nick Herbert. *Quantum Reality*. Anchor Press/Doubleday, Garden City, New York, 1985. [The best nonmathematical introduction to quantum theory that I know.]

[8]Marvin Chester. *Primer of Quantum Mechanics*. John Wiley & Sons, New York, 1987. [The best mathematical introduction to quantum theory that I know.]

wrong. Perhaps there is an evil daemon deceiving my senses into believing in a world. Or perhaps somehow my consciousness is literally all there is, and nothing else exists at all.

As a starting point, anything that can possibly be doubted must initially be set aside. Since all we can know with absolute certainty is our current conscious experience, it is this immediate consciousness, which cannot possibly be doubted, that is the starting point for real science. We may at some future time find ourselves driven back to re-adopting the common sense notion of "world", but then again, perhaps not. For now, we "bracket" the world, and put the matter aside.

Descartes said that there is one thing that cannot be doubted: that I am thinking. This is one thing I can be absolutely certain of. But just what *is* it I am certain of? The frustrating thing about the *Cogito* is that any attempt to formulate it in language seems to introduce at least a tiny bit of doubt. Already, we are introducing the term "I", which is an extrapolation from consciousness. Even terms like "thinking" and "consciousness" introduce at least the possibility that we could be using language incoherently. Properly understood, then, we must separate the absolute certainty of our immediate consciousness from the particular language in which we express this certainty. The certainty of the Cogito is pre-reflective and pre-linguistic. Immediate consciousness does not itself necessarily reflect on the nature of consciousness at all, let alone make linguistic statements about it. It is this immediate conscious experience that is our starting point for science, not some particular expression of it in language.

So what, then, *is* this pre-reflective consciousness? We say it is absolutely certain, but this is not certainty of anything beyond the immediate consciousness of our current state of mind, not even of *this* linguistic expression of that very certainty. Even a fuzzy, unclear thought, in which we cannot clearly understand the object of our perception, is still absolutely certain in this respect. It is only an "unclear" thought with respect to some object *within* consciousness that one is conscious *of*. Yet if we insist on sticking to the absolute certainty of the *Cogito*, we cannot make the move to speaking of this object as something separate from immediate consciousness. It is the object's status in relation to consciousness that is uncertain. Almost any consciousness at all, no matter how uncertain we may be of the objects we are conscious *of*, is equally and absolutely certain in itself. Exceptions may be made here for states of mind where we are only quasi-conscious to begin with, such as certain states of sleep or stupor. However, as we will not be conducting philosophy in such states, for our purposes, we know them only through memory and they will not figure in our explorations. From here on out, then, we will assume that consciousness is always completely certain in itself.

Immediate consciousness is thus our starting point, and we are not yet worried whether it is or is not consciousness *of* anything. In consciousness-of, there is always the possibility of error. Even for something as seemingly certain as "2+2=4", we could just feasibly be mistaken. Perhaps we erred in our calculations. Or perhaps the whole notion of addition is flawed in some way we have not considered. So "2+2=4" does not deserve the label "absolutely certain" in the

same way immediate consciousness does. Yet still, such simple mathematical statements seem to have such clarity compared to almost anything else, that we will want to give them a very high degree of certainty, if not an absolute certainty. For this reason, we will distinguish between certainty in general, which admits of degrees, and the very special kind of certainty that is associated only with immediate consciousness, for which I will adopt Husserl's term "apodictic". This kind of certainty is absolute with respect to consciousness, and cannot be diminished without diminishing consciousness itself. The only kind of "relatively" or "partially apodictic" experience is one that is only quasi-conscious in the first place. Even the vaguest and most confused thoughts of normal waking consciousness are as apodictic as "2+2=4".

With just this, we will get nowhere. We need to make some nonapodictic moves away from immediate consciousness, to make any progress at all. Ideally, we would like these moves to be apodictic, but that may be asking too much. Therefore, we will mark such moves for later reference with an asterisk (*). Each one will require some justification that seems, if not *absolutely* certain, at least very difficult to resist. We must recognize that any such move may need to be changed in the future, and that this may require tearing down and rebuilding everything we have built that comes after. Although we will try to make all such moves as certain as we can, we will find that to retain the absolute certainty of the *Cogito* is impossible if any real progress is to be made. Finally, we note that it is entirely possible, even likely given our fallible natures, that some important unjustified assumptions may miss being asterisked.

The first move away from apodicticity is already contained in the *Cogito*, which says not just that "I think", or even just "thinking", but also "therefore, I am". This is properly thought of, not as a conclusion, an inference drawn *from* immediate consciousness, but purely as a definition. If we view it as an inference, we run the risk of concluding that the *Cogito* validates some pre-existing prejudice about existence, such as material substance or some such. The "being" that is alluded to in the *Cogito* is at this point purely definitional. It points to that which is absolutely certain (consciousness) and gives it a name: "existence" or "being". But we already have a name for it, "consciousness", so why all of a sudden do we need this new term? The answer is that while "existence" does not *necessarily* declare that there is anything beyond apodicticity, it suggests the *possibility*, without taking a stand on the issue. I point to consciousness and say "consciousness", then I say "existence", which I define as being simply consciousness, as something which *may or may not* be more than what is apodictic. So I allow that "being" may end up to be a synonym for "consciousness", but I can assume this no more than I can assume an external world. Either assumption is unjustified. The thing that consciousness is *may or may not* go beyond what is apodictic, but I will use the term "existence" to mean whatever consciousness ultimately is, in either case.

(*1) "Existence" is that which is the ultimate explanation for consciousness, which *may or may not* go beyond the immediacy of consciousness itself.

But even to take this small step, as hard as we tried to restrict its scope, goes beyond our starting point, by making a second-order observation *of* consciousness. Quite aside from the introduction of the term "existence", consciousness has already become consciousness *of* something. Even though we make no claim that existence goes beyond apodicticity, the mere use of the word as something *potentially* different than mere "consciousness" assents to the notion that consciousness is consciousness-of (that it can have an object, namely itself, whatever that is). But we see that to go anywhere from the *Cogito*, we must entertain this notion. This is what generates in us the desire to *explain* consciousness in the first place. A consciousness that moves not even so far as to entertain the *possibility* that there could be an object of consciousness, at the very least consciousness itself, could not ask "why", or "what *is* this thing?". This is the sense of wonder that underpins all science.

While admitting that consciousness can be consciousness-of, we have not yet determined that this is an essential feature of consciousness, nor that it is *of* anything that is beyond its own apodicticity, although we have entertained the possibility. But now that we have admitted that there is such a phenomenon as consciousness-of, there is little if any certainty lost in admitting other instances of consciousness-of. Consciousness of itself, while it provides the sense of wonder, is nothing special with respect to our current move (*1). If I admit that I am conscious *of* my consciousness, I can likewise note that my consciousness is consciousness of a ball, or a rabbit, or whatever. These phenomena appear as things *within* consciousness, but I am remaining uncommitted as to whether there is more to them than what appears immediately and apodictically.

> **(*2)** Consciousness is consciousness-of. The objects of consciousness include consciousness itself (*1), *and* other objects within consciousness.

This forces us to admit that consciousness has a certain internal structure of some sort, and that it is valid to observe and make statements about it, although we do not here claim that this means the things within consciousness extend also outside of it, nor that these things are things at all separable from consciousness itself. However, we *are* making the move of allowing *analysis* of consciousness, without yet understanding *what* such analysis is. This is not a scientific analysis based on mathematical modelling, at least not as yet. Currently, it is simply instinctive. Straightforward observation of our perceptions convinces us that there are objects appearing within consciousness, and such observation we call "analysis".

The word "consciousness", while pointing to the immediate apodictic thing that we perceive, does not fully express the sense of wonder, which by its nature questions whether there is *more* to its own nature, its "existence", than just the apodictic part. We look at objects within consciousness, such as balls, rabbits and consciousness itself, and the sense of wonder instills in

us the desire to know if these objects have a fuller reality than just what appears to us. We are now entering our third movement. Instead of just allowing an analysis of consciousness that somehow involves *parts of* consciousness (the ball, the rabbit, the angry feeling, the memory of roses), we are now exploring the possibility that there may be an analysis that contains consciousness itself *as such a part*. This move is forced on us by the sense of wonder.

We are still not deciding conclusively that consciousness is part of something more than just its own apodictic appearance, but in asking after the possibility, in asking "why?", we are forced to at least assume as a working hypothesis that there is indeed something more, since, although consciousness itself is apodictic, *it is not apodictic to us what consciousness is*. No matter how hard I concentrate, consciousness does not yield up an immediate, apodictic awareness of its own nature. Indeed, consciousness seems inherently unclear to itself. While some objects *within* consciousness seem at least relatively clear, this core clarity quickly fades away into a fringe that cannot be pinned down. Attempt to examine it, and it slips away into nothingness. More than anything, it is this "fringy" nature of consciousness that generates in us the sense of wonder and the conviction that there must be more to consciousness than consciousness itself. We have now begun our search for an explanation of consciousness that goes beyond consciousness.

> **(*3)** Assume for now that there is some explanation for, or truth concerning, our immediate consciousness and what it is. This necessarily goes beyond the immediacy of consciousness, since it is not apodictic to me what my consciousness is.

These first three movements (*1-3) are tightly intertwined, and it is certainly possible to lump them together in a single statement, like the *Cogito*. However, separating them like this will prevent us from falling into the dangerous trap of thinking that the move from consciousness to a larger existence, being somehow "apodictic", is any more certain than "2+2=4".

A Preliminary Analysis of Consciousness

The concept of analysis is crucial. If it is not possible to analyze something in order to understand it, then we can never get past our initial starting point. But we need to know *what* this analysis *is* before we can properly proceed. This very statement, of course, assumes that there is more to analysis than just what appears to consciousness, that analysis itself has a deeper structure that we have not yet necessarily grasped. Our sense of wonder is now forcing us to attempt a naive analysis (the only kind we *can* attempt currently) of analysis itself. So until a more rigorous analysis of analysis is possible, we shall proceed on the basis of plain common sense.

So far, our common sense analysis has shown us that consciousness can be consciousness-

of. In (*2), we have taken this as a defining characteristic of consciousness. In order to bolster this claim, it will help to allow recollection of past apodictic experiences, in order that we might analyze them. But while these past experiences, as recollections, are apodictic, we have no way to be certain they have any past reality as apodictic in their own right, beyond our current recollection of them, or even that there is any past to our consciousness at all beyond the extent of our current consciousness. While immediate consciousness does seem to have some extent in time, this is perhaps only a fraction of a second. While speaking of our memories of the past as valid information about consciousness is, like the experience of the ball-object or rabbit-object, to admit the possibility that there is something more than just what is apodictic, it actually goes further than this. While we suspended further judgement on what exactly the ball-object *is,* while assuming it somehow extends beyond consciousness, we must assume more than just this for our memories, if we are to use them in any special way to analyze consciousness. For if our memory of the past is just another object in consciousness like the ball or rabbit, we are restricted to speaking only of our immediate conscious state when analyzing consciousness, something that slips away into memory before we have had proper time to analyze it. If we could not invoke memories of the past, and trust that they are memories of true apodictic experiences, we would not even be able to appeal to our own sense of personal identity. To analyze consciousness, it is necessary to look at more than just one instance of it. So we reluctantly accept memories of past consciousness as valid:

> **(*4)** Memories of past consciousnesses reveal real instances of apodictic awareness that exist external to immediate consciousness.

Our analysis of consciousness will thus be of consciousness as felt *and* remembered. This is *not* ideal. We want to ground everything in the current moment of consciousness, which while it appears as something in time, cannot guarantee the existence of time in the past or future of immediate awareness. If only we humans were smart enough to hold many multiples examples of objects of consciousness in mind at once, and to analyze them, all without letting any of them fade from consciousness, then perhaps (*4) would be unnecessary. But we need to accept our human limitations, and work with what we have.

In order to properly ground further movements, we must examine various past experiences of consciousness, bringing them to mind once again and if possible reproducing them so as to experience their apodicticity again. Once we have re-experienced them, they immediately fade away, but so long as we can bring them back whenever we want, we accept them, given (*4), as valid examples with which to analyze our current conscious state.

We start by bringing an experience to mind: our perception of a ball. If we just happen upon this perception, then there it is—we did not have to work especially hard to get at it. But if we are remembering a past experience of a ball, getting this perception back into consciousness

may take some doing. We may need to seek out the experience, hunting around for it, which may involve manipulating our perceptions of our body (moving it around in perceptual space until a ball comes to view). Or we may need to focus on our memory of the ball, to bring an image of it back into view. All of these things are perceptions: our memory of the ball before and after we bring the image of it back into view, our perception of the ball when we bring it to view by manipulating our perception of our body, and our perception of our body itself. All are perceptions. We do not succumb here to the prejudice of calling any of them "world" objects. Although we have already assumed that they extend beyond immediate consciousness, this means only that somehow there is a truth to what they are that we are not immediately aware of. This "truth" is not necessarily at all like our everyday notion of an external "world". Therefore, we will call such objects within consciousness, that extend beyond consciousness, "percepts". They are what our consciousness is consciousness-of, i.e., what our perception perceives. Any instance of consciousness is a perceiving act—an act of perception—that perceives objects of consciousness, or percepts (see figure 1). Perception and percept are, of course, inseperable. Together, they constitute a perceiving act. "Perception" refers to the perceiving act from the immediate viewpoint of the perceiver, while "percept" refers to the perceiving act from the viewpoint of the "truth" of the object perceived, whatever that may be. To talk of perception, then, is to take a subjective view. To talk of percepts is to take a more objective view.

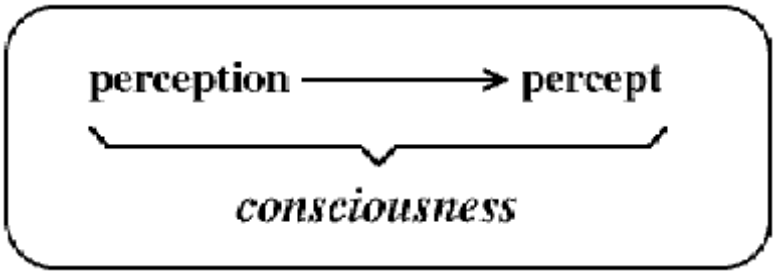

Fig. 1: The Nature of Consciousness: Perceiving a Percept.

Although we came up with this analysis by looking at consciousness, let us remain uncommitted for now as to whether we want *all* perceptions and their percepts to be *conscious* perceptions. We might feasibly discover that our analysis of perception applies in a useful way to a broader range of things than just that which is conscious, and we may find it useful to thus say that consciousness is only one instance of perception. But it is too early to decide on such issues yet.

Although we are not prepared as yet to say much about the objective truth of our percepts, we *can* now proceed to make whatever common-sensical observations about our perceptions and percepts that we want. For instance, we notice that by manipulating the "body" percept we can, with some work, bring a ball percept very strongly into view. Alternatively, by focussing on the memory of past ball percepts, we can also bring a ball percept into view, but this seems to lack

the same strength. Both are, of course, equally apodictic, so we don't mean that the body-ball is more apodictic than the memory-ball. But we seem less certain that we really have a true "ball" percept before us in the case of the memory-ball. This notion of varying degrees of "strength" of percepts, then, presupposes (as so much of our cognition does) that once again there is a truth behind the percept that goes beyond what is apodictic. There is no sense in saying that one ball percept is stronger than another, unless we mean that one is more fully a *true* ball than the other. Strictly on their own terms, they are equally apodictic.

This kind of presupposition infuses so much of our thinking, that the more of these analyses of perception that we do, the harder it becomes to imagine going back on the analysis move (*3). We are constantly presuming that there is truth in our percepts. We presume in the very act of remembering the ball that there is a "true" percept we are trying to bring up, even if we fail to succeed in fully bringing it to view. We feel the body-percept brings the truth even more fully to view, yet the pure apodicticity does not tell us this. Our very perception of the body-percept is *of* something more than just what is in consciousness, and we perceive that this percept brings more of the "true" object into view than the memory manipulations did. Of course, these feelings of truth could all be bogus, but it seems we cannot help but make such presumptions in our very act of perceiving. Even postulating that perhaps these presumptions are bogus presumes some truth to the matter beyond immediate consciousness. The more we examine experiences the more we realize that all experience is perceptual in the sense we have laid out above, and that by its very nature, conscious experience forces us to believe in some kind of truth. Playing the sceptic, we can certainly step back and question this, but not without abstaining from any investigation or analysis of consciousness at all.

Behaviour is the Control of Perception

As strong as the perception of the body-ball is, we still do not feel that the object is *fully* in view. There are still aspects of it essential to its being a ball-percept that are hidden from us. But what are these "hidden" aspects of percepts that are implicit and yet unrevealed in analysis? Are they there in *all* analyses? Is it even possible to bring an object fully into view? These are the issues we will now begin to explore. We will call a perception that more fully brings its percept into view "clearer" than a perception that brings it less into view. So the body-ball is a "clearer" percept than the memory-ball. The "body" thing that is manipulated is probably even less clear, and the "memory" percept is extremely vague indeed.

We see that perception always involves some kind of search through percept-space. Even when we just happened across the ball-percept, it is not as if we did no work at all. We still had to focus our attention onto what was already in view, and pick out the ball-percept from the background context in which it appeared. For the body-ball percept, we searched by manipulating our perception of our body. For the memory-ball, we manipulated our memory. In

all these cases, it was not at all clear exactly *what* we were manipulating. In fact, we do not even necessarily feel when we search our memory that there is something in particular that we are manipulating at all, at least not as clearly as for our body. But even for the body, it is not at all clear what we are doing when we will ourselves to "move arm" or "turn head" and so forth. All we can really say is that we somehow "focus attention" on bringing into view whatever it is we want to bring into view. Our manipulations, whatever they be manipulations *of,* seem to be attempts to achieve certain results in our perceptions. So we thus say that they are manipulations of the percepts within consciousness, although we remain uncommitted as to exactly what goes on that allows us to achieve such manipulations (perhaps the percepts are only manipulated indirectly). All we know is that we feel some kind of indefinable desire within us, that we *want* to perceive something (a "ball" percept perhaps), and that we can "focus attention", in some weird mysterious way, that allows the percept currently in consciousness to change into one that is more like what we want, bringing the desired percept more fully into view. This mysterious manipulation process, achieved by some sort of focussing of attention, we will call our "behaviour". The manipulation of the percept, we will call "control" of the perception (or its percept). Behaviour is the control of perception.[6]

The Possibility Structure of Perception

It seems quite likely that control is a universal feature of perception, for all percepts are at least, if not manipulated, manipulable. Take the ball percept for instance. Why is it that we seem to believe there is more to the percept than just the percept *as perceived*? Careful observation of such percepts reveals much that is "hidden" from immediate experience. The ball, for instance, is perceived as having "other sides" hidden from view. The back of the ball is not apodictically in view, yet our current perception of the ball would not be what it is without the implicit *feeling* that it has a back. While we may not be currently controlling the ball percept in the sense of manipulating it so that we can see the back side, it does not take much reflection to realize that the ball would not be perceived as a ball-percept at all if there was not somehow within our perception the implication of a back side.

For what is this perception of the ball? As an apodictic conscious experience, it necessarily involves more than the ball—everything from the surrounding environment around the ball, to the vague feeling of hunger in my tummy, to the sense of the chair pushing up on me whilst I sit here contemplating the ball. To isolate the ball in this way from the rest of my experience is an analysis, which necessarily includes the possibility of further control, that would bring hidden aspects of the ball into view. It is because I perceive the ball percept *as a ball*, that I feel that I *could* manipulate my body percept so as to see the rear side of it. Likewise, I feel that I could do this by manipulation of my memory-percepts, although we guess from past experience that this will produce a view of the rear side that has less clarity (although we are as yet still in the dark as

to what is behind this difference in clarity).

The memory-manipulation technique we will call “imagination”. We have not at this point established why we feel this is a different kind of control, we simply recognize that there is a kind of manipulation that has a unique feeling about it that characterizes certain types of control. Such imagination-percepts, while usually less clear than body-percepts, seem to be usually easier to control. The body-ball-percept, for instance, we seem to only be able to manipulate within a very narrow range. The appearance of roundness that we see always seems to be there. But when we manipulate the imagination-ball-percept, we can achieve a much wider range of percepts. We can totally transform the ball, morphing it from a round shape into a cube, or into an elephant shape. The price we pay for increased ability to control is a decrease in the strength of the results.

Further reflection reveals something very interesting. Even body-percepts seem to have an element of imagination-control in them. A Necker cube is an interesting example. The reader should now manipulate his/her eyes and head body-percepts so as to bring a “Fig. 2” percept into view. This body-percept will probably appear to you as containing a cube. Its simple uncluttered lines make the imagination component easier to see than in most other body-percepts. It can be manipulated so as to become a cube from (at least) two completely different points of view. With practise, you can readily flip back and forth between both views. The two percepts are clearly different, and yet there is something that is even more clearly seen as fixed and common to both. It is this part that we feel has the strength of a body-percept. The part that we can flip at will seems to be an imaginary component of the core body-percept. Yet any body-percept, no matter how strong its impression, must like the Necker cube be detached from its environment, with its various parts seen in relation to each other in a particular sort of way, before we can even say we are conscious of a percept within consciousness at all.

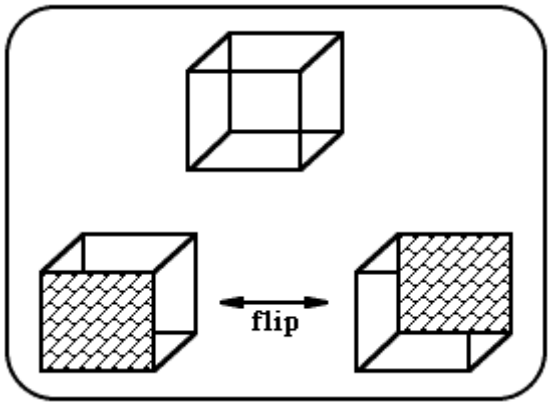

Fig. 2: Necker Cube Reveals the Possibility Structure of Perception.

So is the Necker cube a body-percept or an imagination-percept? Well, it is some of both. We had to use our body to bring it initially into view, but then we used our imagination to manipulate it so as to be a stable percept of some kind. It is harder here than in the ball case to

decide what components of the perception to call "body-manipulated" and what to call "imagination manipulated". The Necker cube can be transformed into all kinds of notably different percepts in imagination (even more than just the two cubes if you work at it), and yet all these different percepts seem united by the strength of the underlying body-percept. We see that to some extent perception is hierarchical, with layers of imaginary perception built on top of other base layers, which may or may not be imaginary. These base layers, on their own terms, contain implicit in them a whole range of possible imaginary extensions. This, we will call the "possibility field" of the perception.

Later on, when the Necker cube is no longer readily in view, we can bring the percept back into view wholly in imagination. The result is not only far weaker, but requires more work. So we see not only that there seems to be an imaginary component to all our percepts, but also that sometimes it is imagination-manipulation that is more difficult than body-manipulation, rather than the other way around.

The Perceptual Possibility Field

In imagination, the possibilities seems endless, whereas body-control seems highly restricted in the kinds of results it can achieve. Yet these results often seem to be much clearer than most imaginary percepts. Are all body-percepts clearer than all imagination-percepts? This we have not yet established. So it is imperative at this juncture to find within percept-space a percept that is truly clear, without bias as to what kind of manipulations this requires. If we succeed, and it turns out to be a body-percept, and we fail to find one that is an imagination-percept, this will lend credibility to the notion that body-percepts are inherently clearer. It might also allow us to finally analyze analysis, which will give our up-to-now naive analysis a firmer grounding. Besides, without *something* that is completely clear, how can we really know that our normal unclear percepts are *of* anything beyond what is apodictic? To show something that clearly has a reality beyond its apodicticity will be to properly justify (*3), not with *absolute* certainty, but with as much certainty as is possible in post-apodictic analysis. If we can make the hidden aspects for at least one percept clear, we can know that the notion of analysis at least makes sense, and can perhaps pin it down. Without this, we are lost.

Concentrating on body-percepts seems to hold little promise. Every time we find something with that "strength" we associate with body percepts, it seems to have an inherent hidden nature that we just cannot fully get at. So next we search imagination-space for a clear percept. Here, while most percepts are even less clear than in body-space, we do find some percepts that just might qualify as "clear", such as "2+2=4". While we are not sure right away if this qualifies, we seem sure that we do have a handle on something that is far clearer than the ball-percept in any of its incarnations. When we try to perceive the ball-percept more clearly, it fades into the fringe of consciousness, without any sense of fully containing the percept within consciousness. But

"2+2=4" is completely different. Here, there is no fading into a fringe. Or, at least, we feel that the core percept, the foreground, can somehow be detached from the fringe or background (see figure 3). We have a strong sense that there is somehow an object here that is completely contained within consciousness.

The "truth" of the "2+2=4" object as something apart from apodicticity appears clearly before us, even though we are still not entirely sure of the exact nature of the object we so clearly see. This seems paradoxical at first, but further examination of other clearly present logico-mathematical objects convinces us that we can *clearly* see that there is a completed truth to a percept without *distinctly* seeing what that truth *is*. The reason is that the truth appears to us as "over-constructed". Yes, we have managed through our control manipulations to construct within our perceptual field a percept without a fringe. Yet this detached core seems itself to have its own inner core and fringe! There is at least a little bit more within the detached core than just the truth of "2+2=4". There are, for instance, notational details that are irrelevant to truth..

The "clear" versus "distinct" distinction is crucial to understanding the nature of what we have and have not gained with the "2+2=4" insight. As something "clear", the percept that lies within a fringed field now somehow seems to stand on its own without the fringe. This is totally unlike the ball percept, for instance, in which we detected an essential incompleteness. Somehow, we seem able to control our perception in such a way as to produce a stable percept that can be seen as a completed object. Like the imaginary morphing of a ball-percept into a cube-percept, or the imaginary flipping of a Necker cube, we can likewise transform this new completed percept into an uncompleted one. We can transform it, so that the other unclear aspects of the perceptual field become "re-attached" to it, so to speak, and suck it back into the vagueness of the fringe.

Examining the truth of "2+2=4", for instance, we can move fluidly from seeing the truth of the statement as self-evident and independent of the rest of the perceptual field, to "re-attaching" it by focussing attention away from the truth of it to the perceptual details such as the particular notation in which it is expressed, the shape of the "+" symbol, et cetera. Now we are seeing something that is not a completed object. We are seeing the body-percept layer underneath the imaginary layer that allows the detachment. Yet, once we learn to see the completed truth, it seems always easy to switch back to the completed view, just like the perceptual control involved in the Necker cube flip.

Further investigation reveals many other perceptions that have this kind of crystal clarity, or completeness. This is completely different from the "apodicticity" of immediate consciousness. Percepts are apodictic without being at all clear, with fuzzy fringes that leave perhaps most of the object completely obscured. Being 100% certain that an object is before you, perhaps incompletely, is not at all the same as being 99% certain that an object is fully before you. Complete clarity, while not apodictic with respect to "truth" or the perceptual object, does

nonetheless generate a high degree of certainty within us. We have a strong feeling that "2+2=4" is a completed truth, even though we recognize some small chance that we could be deluded or in error.

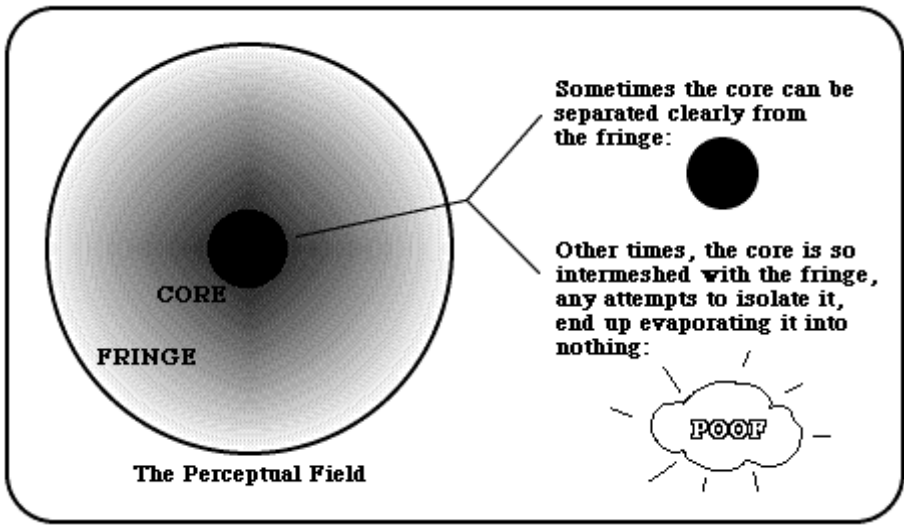

Fig. 3: The Perceptual Field: Clear and Unclear Core Percepts.

Further investigations of perception reveal other completed truths, although the more complex these truths become the less certain we are that we have properly constructed them. After a certain level of complexity is reached, we retain certainty only if we trust our memories and at some further point only if we trust ink on paper. Yet, we remain certain that these mathematical objects are in general there, to an infinite degree of complexity, however uncertain we may be of them individually. For instance, the truth that there are an infinite number of primes seems quite certain when held firmly in one's mind, yet we do not hold all of the primes in our head. Recognition of the truth of certain small, manageable mathematical objects requires a recognition of further mathematical objects beyond what is clearly seen. Seeing the "truth" of "2+2=4" allows us to just as clearly see that there are an infinite number of similar additions that we cannot clearly see at all. It is here where the move from consciousness to consciousness-of is finally clearly seen as revealing an external reality. But that external reality is not (at least not as yet) a material world. It is the world of mathematical objects, some of which are too complex to ever be realized as actual percepts. They are implied in the perceptual possibility field of simpler percepts that we *can* realize. They all seem to be what we might like to call "mathematical", although they do not necessarily involve numbers. Some simply seem to be structures, as in geometrical intuitions, for instance. Any abstract entity that can be precisely and completely described leaves us with this sense of "clarity".

Yet, how very odd that it is at the same time quite unclear what exactly this "clarity" is! The problem thus far is that most of these so-called completed analyses are constructed in various different languages (the language of arithmetic versus the language of geometry, for

instance). In order to determine our criterion for perceptual completeness (i.e., clarity), we need to find a common language that can describe *any* completed percept. If we are successful at finding such a language, we can then attempt to apply it to a description of the more common uncompleted percepts, hoping to be successful at likewise analyzing them (although we will not be guaranteed of success, as it may turn out that the fringy uncompleted percepts require a more powerful language than the mathematical language that describes the simpler completable ones).

While the "truth" of "2+2=4" convinces us we have a completed object in our perceptual field, we see by focussing first on the truth, then on the notational details, then on the truth again, via a Necker-style flipping, that it is not totally clear to us how the clearly present percept differs from other clear objects we can bring into view. If I picture 2 apples being added to a bin containing another 2 apples, and note that there are four apples resulting, my perceptual field as a whole is very different than when I see "2+2=4" written out in arithmetical notation. Focussing on the unclear notational details of the apple-bin-percept (the apples, the bin, and so forth) is quite different that focussing on the unclear details of the "2" symbol and the "+" symbol. Yet I seem to be able to perceive a truth behind both of these that *is the same!* Indeed, here seems to be the key to my feeling that I am experiencing a completed percept. By translating one fringy perception into another fringy perception, via a Necker-style perceptual flipping, I can see something that is the same for both, that due to the translation appears less fringy than either on its own. Only in this case, the flipping does not reveal a core percept that I would likely think of as a body-percept. Instead, it reveals a completely abstract and imaginary core, but one with far more evident objective reality than the body-percept revealed by the Necker flipping. It is fortunate that we did not jump too quickly to conclusions from the Necker cube example, and decide prematurely that such perceptual flipping always reveals a "physical" underlying body-percept. Here is an example where it reveals, much more clearly, an underlying imaginary percept.[9]

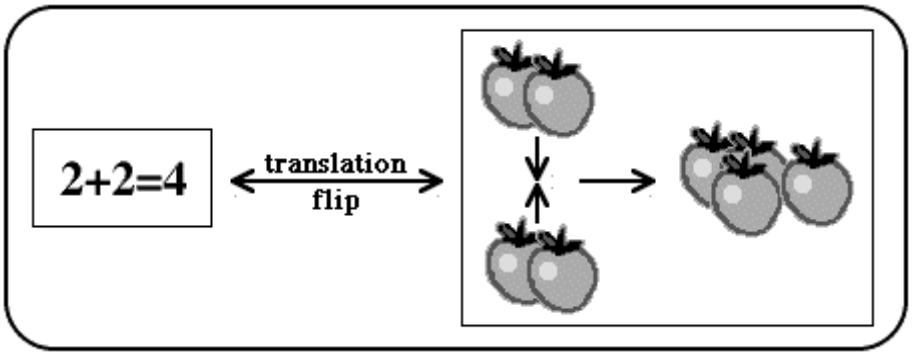

Fig. 4: Flipping between different fringy representations of a clear percept.

[9]The reader should keep in mind that when I say "imaginary percept", I do not mean that it lacks reality outside of consciousness. All I mean is that our perception of it, regardless of its objective status, was achieved through manipulation of that ineffable thing we call "imagination", rather than the ineffable thing we call "body".

I can see now that the perceptual object is something distinct from the apple representation *and* the arithmetic representation. Each representation is "over-constructed", since my construction of the object was necessarily within a fringy perceptual field, and this fringe cannot be 100% detached from the core. At least a little of the fringe always remains. Yet the object itself, the percept, was completely clear (since the "truth" seemed maximally self-evident). Yet it was not completely "distinct" from its construction in alternative languages. The translation from apple-language to arithmetic language, and vice-versa, was helpful, and certainly made the already clear percept *more* distinct, but not totally or maximally distinct. There are always other languages that we have not even thought of in which the truth of "2+2=4" could be expressed. To attain maximal distinctness through perceptual flipping, we would have to translate "2+2=4" into *all possible languages*. This, of course, is impossible in practise, but we can approach this ideal by translating the expression into as many different languages as we can.

Figure 5 shows the basic difference between clear and distinct perceptions using the nonmathematical example of a house-percept. Since, unlike mathematical truths, a house-percept is never completely clear in the first place, figure 5 should be taken with a grain of salt. It is included as a pedagogical device, and not a strict example of a clear perception, for which "2+2=4" is still far superior.

In making this distinction between clarity and distinctness, we seem to have hit upon an even better understanding of the analytic nature of perception. We saw earlier in our analyzing of the ball-percept that there was inherent in the percept a range of possibilities not immediately present to consciousness. Not all perceptions are part of this "possibility field" that surrounds the core of the percept. For instance, the perception Abe Lincoln had just before being shot cannot be part of my possibility field for my ball-percept, since I cannot through mental control flip myself into a Lincoln perceptual state.

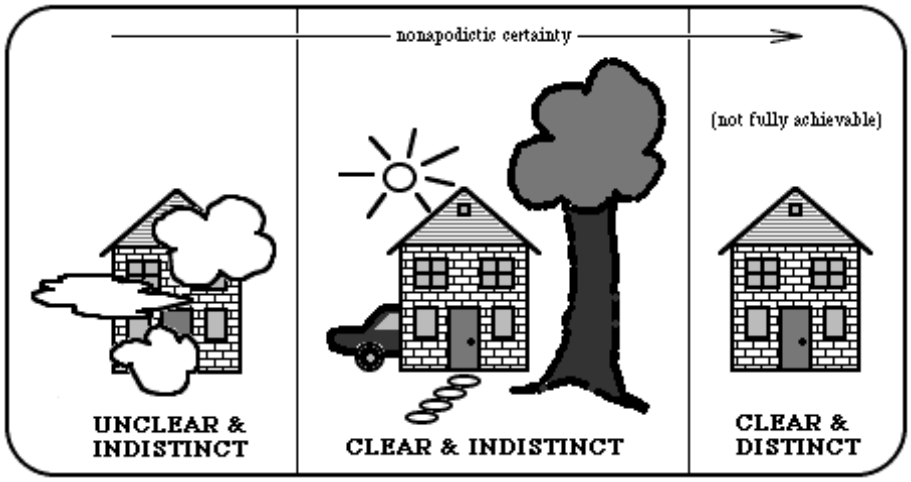

Fig. 5: Three apodictic perceptions, with increasing house-certainty.

On the other hand, maybe I *can* flip myself into Lincoln! That I cannot seems thoroughly

reasonable, but is actually an assumption that we cannot fully justify here, one that will be necessary in what follows. Personal identity, we presume, constitutes a particular flow of consciousness. Perceptions in my future are ones that remember being me in the past. Perceptions in the past are ones that I can remember being. None of Abe Lincoln's perceptions can remember being me, nor can I remember being them. Exactly how it is that I can unite various perceptions into a personal identity is not yet clear, but for now, we will have to assume that I *do* have an identity, and that this includes some possible perceptions and not others.

> **(*5)** My personal identity includes only some possible perceptions and not others. All are somehow implied in the possibility field of my current consciousness. Perceptions that remember being me are (at least) my future self. Perceptions that I remember being are (at least) my past self.

So the possibilities in the perceptual field are only a subset of all possible perceptions. Or, to be more precise, some are at least to be included more fully than others. Even if Abe Lincoln's consciousness does make it into my possibility field at some point, it will necessarily have to be far indeed from the core if my own identity is to be retained. For the possibilities in the field are by no means evenly distributed. Some of the possibilities are much easier to control for than others, and there are certain kinds of percepts that are much more common near the core than others (shape percepts are probably more common near the core of the possibility field for the ball than for the Necker cube). Moreover, we can split the perceptual field into various parts, according to our understanding of personal identity. Some perceptions are implicit in the field *as* future perceptions, some as past, others are more or less explicitly involved in the current percepts under control right now. Others are merely implied as mathematical possibilities that are unattainable in practise.

All aspects of the field are tightly constrained by two factors: (1) mathematical possibility and (2) personal identity. There is much room here for debate as to what should be included as part of this "field". No clear method has yet been given for specifying the full field of possibilities. Should *all* the mathematical truths implied by "2+2=4" be included, even if they are too complex to *actually* achieve a perception of? Taken to an extreme, all possibilities might thus be included in every perceptual field. How then could we distinguish between different perceptions, or different individuals? The answer is that, however great a scope is given to the "field", the possibilities included have a definite distribution, with some closer to the core than others. In other words, the field must not be defined simply as a set of possible percepts, but a set of percepts and possible percepts with a definite distribution that divides into four main parts:

(1) a *core* which is the percept as perceived in consciousness (e.g., ball-percept),
(2) a *fringe* of consciousness into which the core gradually fades (e.g., background noise),
(3) a *possibility field* of attainable percepts within the fringe (e.g., other hidden views),
(4) a *possibility field* of unattainable percepts in the outer fringe (e.g., all possible shapes).

These are not necessarily so neatly separable from each other, but our analysis so far has seen a need for all four basic kinds of percepts. The term "perceptual field" could be used narrowly to mean 1-2, in which case the hidden views of the ball would not be included unless we were explicitly conscious of them. More broadly, the term could mean 1-3, in which case the hidden views are included, but only if it would actually be possible in practise to control so as to bring them into view. The full field considered as 1-4 includes all possibilities implied by a full analysis of the core percept. Even so, these percepts are to be considered as laid out in a field, so that some are nearer the core than others. Exactly how this distribution is constructed will depend on a fuller understanding of perception that we just do not have as yet. For the purposes of this essay, we will call field 1-3 the "personal identity field", since it is within this scope that Abe Lincoln's perceptions are definitely excluded from my field.

Exactly how broad we want the "field" under consideration to be, and exactly how we want to construct the possibility distributions, will depend on the nature of the manipulation we are attempting. Different kinds of control will yield different distributions of possible percepts. The translation of "2+2=4" into different percepts with the same core is an attempt to detach the core from the fringy field. It is thus just another example of perceptual control, where "truth" is the stable thing being controlled for (rather than a stable visual image, etc.). All perceptual control seeks stability and thus some kind of detachment of the core is being sought. Clear percepts are simply those in which we feel that the detachment is complete, rather than merely relative.

An Analysis of Analysis: The Nature of Mathematics

Before we go further, however, we need a clearer picture of analysis itself. We have seen that it involves a recognition of what is common between different percepts. This is needed in order to focus attention on the core and bring it forth, like the Necker cube interpretation, as foreground against the fringier background. Such analysis of a percept necessarily involves implicit recognition that the field and fringe of the percept brings many other possibilities into play (some more strongly than others) than just what is apodictically before consciousness. The process of control is our ability, through a mysterious act of will, to somehow "navigate" our way through this perceptual possibility field (where subfield 1-3 is the part that is in practise navigable). Although we sense that there is some definite structure to the probability distribution of this field, we are still in the dark as to the exact nature of it. For this analysis we will need a clearer and more distinct picture of analysis itself. But will analysis itself be amenable to analysis, with the total clarity and relative distinctness we seek? All we can do is give it a shot and see how it turns out.

The ability to translate from various different representations seems crucial to the objective nature of analysis. So far, we have seen that analysis has a Platonic character to it. We seem, in translating the percept, to capture more of its distinct objective essence. To do this, we needed to

be able to recognize what was common between the different perceptual fields. Plato would say that we were accessing the "form" of which the different perceptions were mere instances. So we will adopt this language and call whatever it is that is common between different perceptions the "form". A true form has not been completely described so long as the percept remains unclear. A clear percept yields a completely constructed, but over-constructed, percept. We will now try to analyze the currently vague and unclear "form" percept into something clearer.

What we are essentially attempting in the following is to construct a universal language. This language will fall short, of course, since it must itself be expressed in a particular language, but if we can capture in one language some of what is common between them all, we can at least express our clear percepts in a single language that provides near-maximum distinctness, and this will prevent our having to constantly translate between all possible languages we can think of, just to make a percept more distinct. But just as we could never translate between *all* possible languages, if we do find a language of near-maximum distinctness, this language will, like our multiple translations, only provide a relative, not total or absolute, distinctness.

Church proposed a formalism called the λ-calculus that has the universality we are looking for. It is simple and can express any mathematical truth that can be expressed. Let us start with a specific example of a form from everyday life, say a health insurance form that needs to be filled in with your name and address and the like. A single health insurance form can be filled in with many different combinations of names and addresses. When John Smith and Jane Doe fill in the same application form, they end up with two different particular instantiations of the same general form. To instantiate the general form into a particular filled-in form, they fill in the blanks. As in:

Name: _____ → Name: John Doe
where "→" represents the "filling in" process

Each blank is a "variable". We will represent the form as an "S-expression", which stands for "symbolic expression". This is just a list (or ordered set) of symbols contained within parentheses, such as (a b c). The elements of the list can themselves be further lists, as in (a b (c d) e). This is a "concrete S-expression", since it does not yet have any "blanks". The notion of a form is not yet there. Although many different "structures" can be built with this concrete notation, it fails as a universal language, since there are many things that can be brought clearly into our perceptual view that nonetheless escape analysis as concrete S-expressions. "2+2=4" can only be so expressed if we accept a lack of clarity. For instance, if we choose the S-expression (+ 2 2) to represent "2+2", we have only a list of arbitrary symbols. They just sit there. None of the *process* involved in computing the sum of 2 plus 2 is captured here. Concrete S-expressions form an inadequate language in which to represent clear percepts.

We need to add the blanks, so that we can fill them in and truly have a form. We call this

abstract S-expression a "λ-expression". It is written as follows:

(λ**x.P**)
where **x** and **P** stand for S-expressions (perhaps concrete, perhaps abstract).[10]

In the above form, **P** is an S-expression, which may (or may not) contain instances of the variable **x**. The **x**'s are "filled-in" when this form is given another S-expression, call it **I**, to "fill in the blank" wherever an **x** appears in **P**, producing a new filled-in S-expression, call it **O**. This filling-in is called "evaluating" the form, or sometimes "applying", "simplifying" or "reducing":

((λ**x.P**)**I**) —evaluate→ **O**

For example, if Jane Doe were filling in an insurance form, the part that asks for her name could be formalized (i.e., constructed clearly as a form) as follows:

((λ**____ .** "**Name: ____**") (**Jane Doe**)) —evaluate→ ("**Name:** (**Jane Doe**)")

The "λ" part of the form, to the left of the period, simply defines the "blank" to be filled in. To the right of the period is the expression containing the "blanks" that actually get filled in with whatever further expression the form is "applied to", appearing to the right of the λ-expression.

λ-forms get much more interesting when we allow Jane to fill in the "Name" blank with yet another form! Perhaps she wants the application form to apply to the entire Doe family, so in the "Name" blank, she enters "____ Doe".

((λ**____."Name: ____"**) (λ**____.(____ Doe**)))
—evaluate→ (**"Name:** (λ**____.(____ Doe**))**")**

Notice that although the form evaluates to an expression that contains yet another form, no further evaluation is possible, since the internal form is not being applied to anything—nothing is given to fill in the blank. Furthermore, this final expression cannot itself be applied to anything, since it is just a regular concrete S-expression *not* a λ-expression. However, it does *contain* a λ-expression. This form within a form could be applied internally, within the S-expression, to either "Jane", or any of the first names in Jane's family, such as that of her husband John. Evaluation would then require two steps before giving a result that cannot be further evaluated:

((λ**____."Name: ____"**) ((λ**____.(____ Doe**)) **Billy**))

---

[10]The bold-face characters are place-holders for S-expressions, but are not themselves S-expressions. They are a way of talking about S-expressions, and are called, along with the S-expressions containing them, "meta-expressions," or "M-expressions." Most so-called S-expressions you see in print are actually M-expressions. An M-expression that stands for a λ-expression is called a "function."

—evaluate→ ((λ____.”**Name: ____**”) (**Billy Doe**))
—evaluate→ (**“Name:** (**Billy Doe**)**”**)

Again, we end up with a final S-expression that is not a λ-expression: not abstract, but concrete. From a Platonic perspective, this concrete object lacks objective reality; it is not anything in itself. It is the “filling-in” of the blanks, the evaluation, that is “true”, not the end result, which has no truth apart from our conscious perceptual field. Concrete S-expressions are non-entities apart from their particular perceptual field, fringe and all. In other words, they do not exist in themselves, apart from our interpretation of them in a particular perceptual “basis language”. A concrete S-expression cannot be coherently separated from the entire context of the scratches and markings on paper, my seeing of those markings and my understanding them in a certain way. λ-expressions, however, *can* be so separated (although this detachment of core from field is necessarily inadequate and falls short of total distinctness, as we have already established).

If a λ-form is complicated enough, its evaluation can require many more than just two steps before halting like this. In fact, it might *never* halt. Note that “never halting” does *not* mean “requires an infinite number of steps to halt”. The latter is a nonsensical statement—there *are* an infinite number of steps in the evaluation, so the λ-form *never* halts. There is no final result that we could clearly perceive if only we could evaluate the form for an infinite amount of time. An infinite evaluation simply *has no* final result.[11] The full infinite running of the form can be considered as part of the perceptual field 1-4, but never 1-3, since we cannot in practise perceive all of it.

It is easy to see that the process of substitution at each evaluation step is analogous to physical time, and thus we could call it “λ-time”. We can also see that any percept that can be clearly detached from the fringy field thus necessarily has a temporal quality. Static percepts are fictions, necessarily unclear and nondetachable from their perceptual fields without introducing a temporal component. If we later discover that the entire perceptual field can be analyzed as a form, we will be able to declare the same for all conscious perception: that it is necessarily temporal. But since we are not yet justified in concluding this, we will simply state for now that it is consistent with our observation of perception in general, which does indeed seem, even when very unclear, to always contain a temporal character. Try to observe, by focus of attention, some typical unclear percept, such as a visual image. This is very difficult, as the percept keeps changing in our very attempt to observe it, as it moves within its possibility field. But getting as close as we can get to pinning it down and detaching it from the background field, we find it seems irreducibly temporal in quality.

---

[11] Although it is possible the expression could approach a particular S-expression in the limit. However, it takes *another* λ-expression referring to the nonhalting one to establish this. *In itself* the λ-expression just produces an infinite sequence of new expressions.

The result of each step of a λ-evaluation depends on the previous result, so evaluation can be thought of, not just as abstraction, but also as a model of causation or mechanism. So we see that "abstraction", "causation" and "mechanism" are just different words for the same thing. You may have guessed by now that the λ-calculus can also be viewed as a simple and elegant model of computing (in fact, the popular computer programming language LISP is a version of the λ-calculus[12]):

((λx.P)I) → O
P is the "program"
I is the "input" or "environment"
O is the "output" or "value" of the computation.

Thus far, this is still just a candidate for a universal language. How do we know that *any* clear percept can be represented in this basis language? Once again, we must fall short of full justification. All we can do is search through percept-space for clear percepts, and see if we can translate them into this language, thus achieving relative distinctness (and perhaps also translating into as many further languages as we can to make the percept even more distinct). However, much experience in the field of information science has convinced many that this simple model *can* account for anything that can clearly be perceived, i.e., all of mathematics and logic. The reader, to fully appreciate this, must spend some time searching percept-space. Some experience programming computers helps considerably to see this. The "Church-Turing Thesis" expresses the idea, which is commonly taken as a truth in information science, although it has never been proven in any of its incarnations.

(*6) The Church-Turing Thesis: all possible clear percepts (i.e., all mathematical truths) can be expressed as λ-expressions (i.e., recursive functions).[13]

It must again be re-iterated that it does not follow that *unclear* percepts can also be captured in the λ-calculus. This is a vitally important issue to resolve (although one we will not settle on in this essay). Those who believe that *all* percepts, clear and unclear, can be captured in such a formalism, are sometimes said to adopt the "strong hypothesis of AI (artificial intelligence)", that an appropriately programmed computer (i.e., an appropriately constructed λ-form) could in principle be conscious.

By "in principle", it is meant that although we ourselves may lack the resources to make our own percepts clear and distinct, they *are* as thing-in-themselves, in truth, analyzable forms.

---

[12]In fact, I have used the terminology of LISP in this paper. The actual terms used vary somewhat from one version of the λ-calculus to another.

[13]Church and Turing, of course, did not formulate their version in the phenomenological terms we are using. They talked of things that would "naturally" be considered as "computable". However, many different versions (some ontological, some epistemological, some cognitive) of the thesis are held by various schools.

In other words, they are in themselves distinct things. Perhaps they could be made clear to larger-brained creatures, superbeings, that were so much smarter than us that they *did* have the resources to fully analyze us and make our percepts clear *to them*. Those who oppose the strong AI view, however, believe that, while our unclear percepts may have a formal component, the universal language we have constructed above is not sufficient to describe the full range of our perceptions—our unclear percepts, complete with field and fringe, would escape analysis even by the most intellectually gifted superbeings. This issue we cannot settle yet, and we will not settle in this essay. However, until such time as we know how to look for a better universal language, or have specific reason to believe our current language is incomplete, our only sensible research path is to continue to try and analyze perception with the formal tools at hand, and see how far we can go towards making all our perceptions clear and distinct.

> **(*7)** The Strong AI Hypothesis: all perception is clearly perceivable, at least in principle (i.e., it consists of clear percepts, expressible as λ-expressions).

It is important to note, however, that until we have achieved clarity, we cannot jump the gun and think that our formal structures properly reflect the structure of the things-themselves. As we build tentative formal structures that capture *some* of what we observe in our perceptions, we certainly hope that these formal structures have something to say about the things themselves, but so long as the percepts we are analyzing remain to some extent unclear, we must proceed with great caution, and expect that further clarity may completely change the nature of what we think we have before us.

Our formal language, the λ-calculus, can itself be made more distinct by translating it into various basis languages. A search through book-percept-space via body manipulation reveals many different formulations of universal language, the equivalence of which has been demonstrated with great clarity.[14] One is called "Turing machines". Another is called "first-order predicate calculus plus set theory". Another is called "cellular automata", another the "NAND-gate". Yet another is called "English" and still another "German". These last two are not very efficient at expressing clear and distinct percepts, but they are adequate to the task.

We discover that whatever can be expressed with total clarity in any of these languages can be made more distinct by translating it into as many of the others as possible. All such languages are called "Turing-equivalent". We also find that the λ-calculus is the simplest and most elegant single language that we have yet found, and thus best captures what is common to the entire equivalence class of languages. It best formalizes formality itself, and thus is the single language most capable of yielding distinctness, although no more able to give us absolute distinctness than

---

[14]George S. Boolos and Richard C. Jeffrey. *Computability and Logic* (3rd Ed.). Cambridge University Press, Cambridge, 1974, 1989. [An excellent technical introduction.]

we are able to translate between an infinite number of languages.

This class of languages is sometimes called the set of recursive functions. This gives us the full power of arithmetic and number theory and the rest of mathematics. Thus I will proceed to use mathematical formalism freely without necessarily expressing everything explicitly in the λ-calculus, which is sometimes too awkward. Anything that can be precisely described mathematically, and programmed into a computer, can be expressed as a recursive function, and will thus be grounded in our universal language. Now that we have analyzed analysis, we are prepared to more properly analyse the perceptual field.

Hierarchical Perceptual Control Theory

The notion of control of the perceptual field in order to achieve a stable percept involves the comparison of a percept, somehow extracted from the given perceptual field, with a "reference" signal, which is the goal to which we seek to bring the percept, roughly corresponding to "desire". This whole process takes some time to occur, naturally, if it is to be achieved via λ-substitution. Consciousness behaves so as to control its perceptual field *for* a particular reference percept, which may simply be the presence of a stable image, or may be more specific such as the presence of a ball-percept (and as long as we fail to bring a ball into view, we continue searching). The act of focussing attention so as to bring the percept closer to its referent is "behaviour". Behaviour, said W.T. Powers, is the control of perception.

Figure 6 shows a simple schematic of the simplest type of control system, which combines various elements of the preexisting perceptual field to produce a new percept, according to the "basis" specified by the PERCEPT() function. This is compared to the referent by the COMPARE() function, the result of which is used to control the percept via the OUTPUT() function, which transforms the result of the comparison into an action that will (hopefully) bring the new transformed perceptual field more into line with the referent.

If the PERCEPT() and OUTPUT() operators are taken as simple sums and COMPARE() is taken as a subtract function, we have the simplest possible feedback control system. I am not claiming here of course, that perceptual control is really built like this, just that this is a recursive formalization that captures what we have been able so far to analyze concerning perception. Future discoveries of the structure of our perceptual field may force us to fill in more details, make the situation more complicated, or even force us to redesign our model completely.

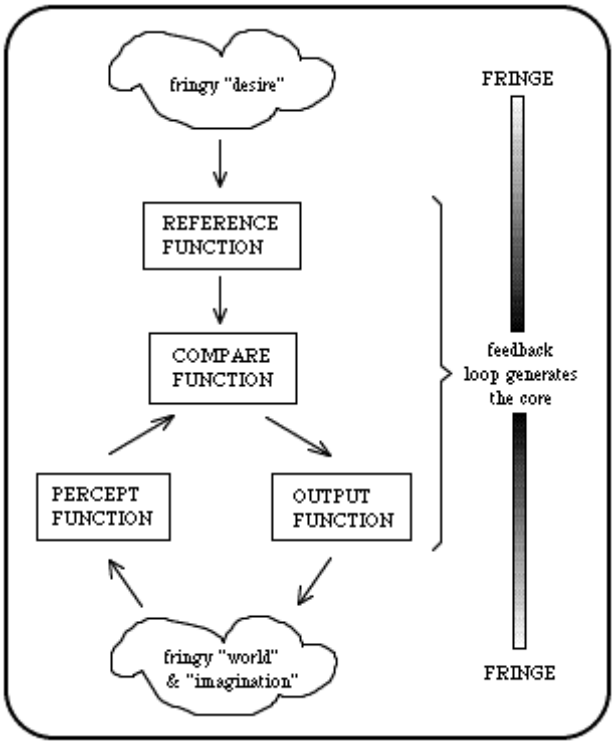

Fig. 6: Perceptual Field Generation.

Some reflection will reveal that there is a hierarchical structure to perception, which we have already noted in outline. Perhaps the thing-in-itself is not literally formalisable as a strict hierarchy, but there is some kind of hierarchy there after a fashion. Take for example, the act of driving to work in your car. You are continually controlling for perceiving yourself near the centre of the road by acting on the steering wheel percept. But why are you doing that? Examine yourself very carefully, and you will discover that you are in fact controlling for getting to work, only on a somewhat longer time scale. On an even longer time scale, you are controlling for making money, and ultimately for being happy. Some of these references may only be on the fringe of your consciousness, even nonconscious, in spite of the fact that they are influencing your every move.

The simplest model we can build with recursive functions would look something like figure 7. The referent at each lower level is supplied by the behaviour of the higher level. The referent for the staying-on-the-road subsystem is supplied by the behaviour of the getting-to-work subsystem, for instance. The perceptual input for each higher level system (the input to its PERCEPT() function) is taken from the percepts of all the lower level subsystems.

At all levels in the hierarchy we are acting to achieve a stable percept, according to the basis language in which our perception is expressed, determined by PERCEPT(), COMPARE() and OUTPUT(). This is the feedback loop. We could feasibly have all kinds of different feedback loops within us, some perhaps in conflict with others. This would certainly explain certain feelings of tension that we feel all the time (such as “I want to go to work, and make

money, but at the same time I want to stay in bed and sleep").

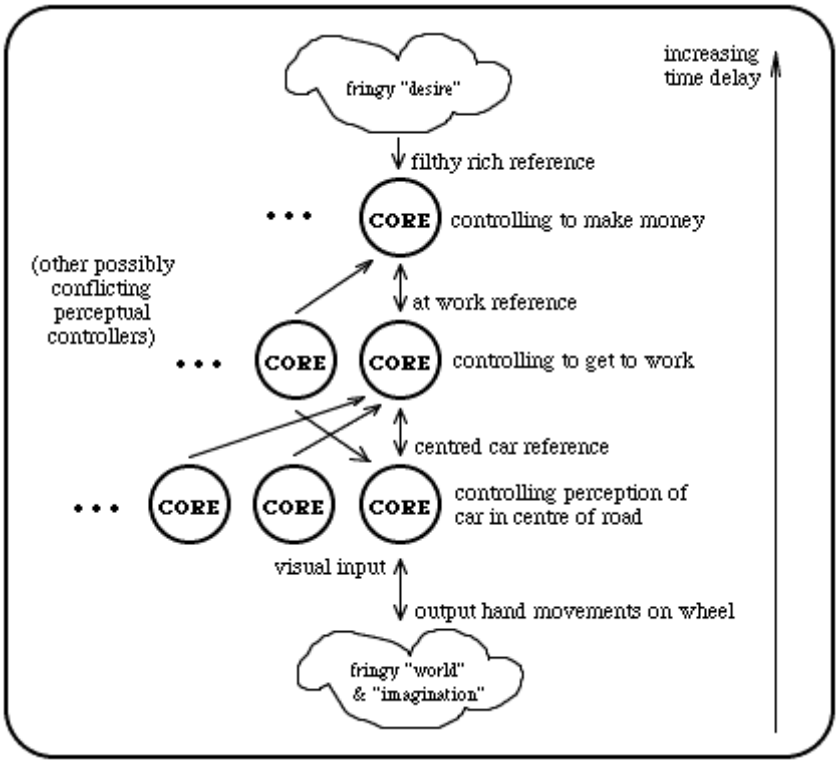

Fig. 7: Hierarchical Perceptual Control: Driving to Work.

Figure 8 shows a graph from a computer simulation of a simple control system controlling its percept, bringing it in line, over time, with the referent. When there is a disturbance to the referent, the system acts to bring perception back to a stable state. This kind of behaviour is universal in all things that we call "living", not just those we call "conscious". Earlier we said that the terms "perception" and "percept" were being kept open for possibly wider applicability than just conscious perception. Here we see an appropriate broader definition. The simple system depicted in figure 8 is a single comparator that does a simple subtract of a scalar percept from a scalar referent. As such, it is no doubt not even remotely conscious, but it *is* a control system with an identifiable percept. If we were given a larger system in which this controller was embedded, it would be possible to isolate the controller clearly from the rest of the system as something that perceives, even if we were not told ahead of time what was "organism" and what was "environment", and even though consciousness is not really even an issue at such a simplistic level (we would simply look for stabilities that maintain themselves in the face of disturbances).

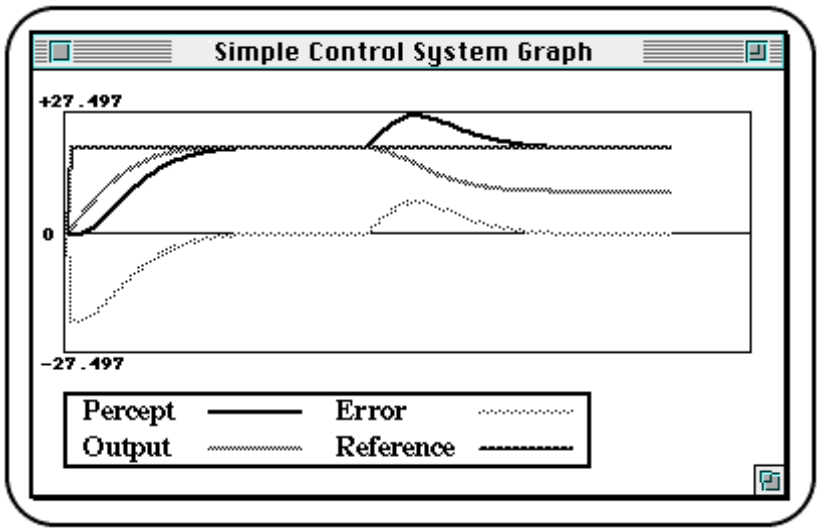

Fig. 8: A Simple System Controls Its Perception Against Disturbance.

It is entirely feasible that our own consciousness arises from the unconscious behaviour of many such mini-controllers working in parallel, and that as we move out in the hierarchy from the conscious controllers in the core, to the fringier controllers, consciousness just gradually fades into nonconscious perception, without there being a firm line between the two. This is a reasonable hypothesis, as it builds our perceptions out of lower level entities that work essentially the same as the ones we built from what we actually know about our own conscious perceptions. But that does *not* mean that the hypothesis is necessarily true, and since we will not be developing or justifying our theory of cognition any further in these pages, I will leave it at that for now.

The Distribution of the Perceptual Field: Quantum Mechanics

We call the different languages in which a stable perceptual object can be expressed different "bases" for the percept, determined by the exact layout of the control system. Returning to the "apple-language" versus the "arithmetic" language for the "two-plus-two" percept, these two bases are perceptually quite different, and yet both have been used to construct the same stable, fully detached core percept.[15] Note that the basis for the control of one percept consists of the percepts from all the lower level control systems (until perhaps we get to the "lowest level" of control, if there is one).

When we analyze a percept through the use of a controlled manipulation, we move from one percept to another, via a manipulation in which is implicit some language we have chosen as a basis on which to form a stable percept. This sometimes appears as a forced act of will, especially over longer time scales, but appears to work more automatically for perception on the faster time scales, as in image perception. Here, while the work is done more automatically, the

[15]Recall that "fully detached" means that what we detached we completely detached, not that we managed to detach exclusively what we were after without dragging anything else in with it (the percept is clear but only relatively distinct).

percept still involves control of a stable percept, forming a foreground out of the background. If you examine yourself carefully, you will see that you *are* doing some work, although it is relatively effortless. The basis language, as part of the perceptual field, is perhaps readily accessible, perhaps not. It can thus at least potentially be brought to the foreground as a percept itself, although perhaps in practise only by superbeings capable of fully understanding us. Nonetheless, the basis language and the percept-object are both possible percepts (*7).

We will represent the basis of a percept, itself a percept, within the "< |" symbols, called a "bra" for reasons that have nothing to do with women's underwear, and which will become clear shortly. In the current example we have two different basis languages for the two-plus-two percept:

< apples | < arithmetic |

The thing-in-itself that is being perceived (and we are justified in calling it a distinct thing-in-itself to the extent that we have been successful at detaching it) is represented in the complimentary symbols "| >", called a "ket". I will use "two-plus-two" to represent the thing-in-itself for the "2+2=4" example, although a full expression in the λ-calculus would be a better choice:

| two-plus-two >

The ket-percept, if we have achieved a clear analysis, should be represented as clearly and distinctly as possible. Thus, we call it the thing-in-itself. The bra-percept, however, can be very unclear, as it involves the perceptual field in which the thing-in-itself is represented. Very often, however, the ket will also be unclear, but it is important to remember that this is the part that we are always attempting to make as clear and distinct as possible.

The entire perceptual field, fringe and all, of any one particular expression of the ket-percept in some basis language is given by combining the complementary brackets, to form a bra-ket, or bracket:

< apples | two-plus-two > , <arithmetic | two-plus-two>, *or in general:*
< perception | percept >

What is placed within both the bra and the ket are *both* percepts, and are not fundamentally different kinds of things. The thing-in-itself is put in the ket. The language in which it is expressed is put in the bra. Any particular perception of the thing-itself is necessarily "bra-ketted", i.e., expressed in some particular language, and thus not wholly distinct apart from it. Our perception of two-plus-two cannot be wholly separated from our perception of it in a particular language. We thus cannot ever fully and adequately know the perceptual object as a thing-in-itself, no matter how clearly it may stand before us, since no matter how fully it is

constructed in some basis language, the necessity to choose a limited number of basis languages will always mean it is over-constructed. The λ-calculus, while a wonderfully elegant language, is in the end just another language with no absolute metaphysical status. The two-plus-two percept, as a distinct object, cannot be expressed *for us* within consciousness in all possible basis languages. We can never know |two-plus-two> directly, without a bra. Nonetheless, we speak as if we somehow have accomplished this when we achieve a Turing-equivalent expression of the percept, and so we say that the role of the ket is to contain what is known:

Quantum Rule #1: Put what you know in the ket. Put your unclear perceptions in the bra.

Since the ket-percept is clear (i.e., complete), it is recursive, and can be expressed functionally. The bra-ket notation can thus be re-expressed functionally as follows:

< perception | percept >	*(bra-ket notation)*
percept(perception)	*(functional notation)*
(λ percept) (perception)	*(lambda notation)*

The functional notation, of course, is only one arbitrary way of viewing a form, for which the argument to the function (the perception) is not a fundamentally different kind of thing from the thing-in-itself (the percept). But how can we pass the "perception" as an argument to the "percept" function if it is, unlike the percept, unclear? The answer is that, given the strong AI postulate (*7), we can do it *in principle,* since the entire perceptual field is formalisable, at least by super-beings, although probably not by us (so we will continue to talk of the bra-perception as unclear). In reality, given the limits of our resources, both the bra *and* the ket will usually remain less than completely clear. Another way to think of the distinction is that the "bra" represents a more fundamentally subjective perception than the "ket". Think of the bra as "my perception seen through my own eyes", and the ket as "my perception seen through the eyes of the super-beings".

The computation that is performed by the function call percept(perception) will, if fully formalized, actually produce the entire conscious perceptual field that is consciousness *of* the specified percept. This is not a static object that the computation somehow produces as an "output value". Rather, the very object that is consciousness is itself a process, a necessarily temporal thing. It is the very running of the abstract computer program that is the perceptual field, not some static end result. The bra-ket <apples | two-plus-two> (or function call two-plus-two(apples)) will produce the entire perceptual field that is consciousness of the objective truth of two-plus-two, via focussing attention in imagination on bins with apples in them. The two-plus-two() form could also be instantiated by passing it a different perceptual basis, as in two-plus-two(arithmetic). This computation produces a different perceptual field, but one that is implicit in the possibility field of the other. What the two perceptual experiences have in

common is the clarity of the "2+2=4" truth.[16]

So two-plus-two(apples) yields our conscious apple-adding experience. The apple-basis, as mentioned earlier, can be detached as a percept itself, which we (perhaps) use to manipulate or transform the "2+2=4" arithmetical percept. We think of this as control within imagination. When it happens more automatically, without conscious willing or focus of attention, we think of it as control via the body and are more apt to call it an "observation" than willful control. An example would be the observance of a ball by directing our eyes towards it. Say there is a table on the other side of the room, possibly with a ball under it. We look up from the page, directing our eyes in that direction. There are many possible percepts we could perceive by doing this. The field of possibilities becomes actualized as we move our eyes to the table. We might, for instance, see the ball on top of the table, or we might see it underneath. Both are perceptual possibilities, and both are easily brought to view within imagination. If we check for the ball's location in this fashion, we could describe the perceptual field as ball-percept(ball-on-table), or <ball-on-table | ball-percept>. We are focussing our attention on the ball in such a way as to see it against a background, as lying at some point in space before us.

<ball-under-table | ball-percept> would be another possible outcome of the same body-manipulation. Both "ball-under-table" and "ball-on-table" are given the role of basis languages. This is true at least from one point of view, in that the thing-in-itself, the ball-percept, can be seen either as under the table or as on the table, so these can be viewed as basis languages via which to perceive the ball. However, we can step further back and observe the manipulation itself and see the higher level language that includes both ball-on-table and ball-under-table as different kinds of "ball-seeing", call it the "look-ball" operator. There are infinitely many ways, in fact, to analyze the perceptual field of possibilities into languages within languages of all different sorts (especially within imagination). "look-ball" is the observation we make, and is the language we have chosen that determines the kind of control we use in order to achieve a stable percept.

In order to include in our representation both this basis that we are calling the "look-ball operator" *and* the more specific ball-under-table basis that we already placed in a bra, we here introduce the general idea of an "operator", which will be defined as an act of manipulation that transforms a percept in its possibility field. If we have but one operator, it is just equivalent to the basis, or bra. If we have one operator that acts on a given basis, we call it an operator *on* the given basis language. Thus, for the ball percept, the "look-ball" basis can be viewed as an operator that we use to focus on the ball-percept, producing the (perhaps unclear) ball-on-table observation. We enclose an operator in vertical bars, "| |", so that it can be placed with the bra-

[16] To make the entire arithmetic version clear would mean performing *((*λ two-plus-two)(arithmetic))(language*X*), where language*X* is some language that allows us to represent our entire perceptual field within itself. (It is likely that language*X* cannot exist for us, although the computation could perhaps be carried out by the superbeings).

ket:

ball-percept(look-ball(ball-on-table)) *(functional notation)*
<ball-on-table | look-ball | ball-percept > *(bra-ket notation)*

The above bra-ket expression could be read as "the event of controlling a stable perception of a ball on the table *given* that I am controlling for perceiving the ball by looking at the table, *given* the thing-in-itself that I am observing." The vertical lines can thus be read as "given" or "given that".

Figure 9 shows two different operators that we will use in our example: look-ball and look-sky, which involves looking up from the page and out the window at the sky, instead of at the table. Two possible perceptual results are shown for the look-ball operator, and one for the look-sky operator. In reality, there are many more possible results. Note that the possibilities within the perceptual field are necessarily restricted here by our choice of operator/language. We are talking about all possible perception that could result from this action. This is necessarily restricted by what is (1) mathematically possible and (2) by our personal identity. The consciousnesses that perceives these "results" must thus in some way "remember" being me looking down at the page, or at least this memory must be accessible from within the perceptual field (1-3). The precise criterion to be applied here would require a more complete theory of consciousness, but the basic idea is simple enough: which mathematically possible perceptions count as possible within this context is determined by my personal identity.

There are, of course, far more possibilities than are shown in the diagram, although we will find it convenient to talk as if these were the only possibilities, as a pedagogical tool only. Please take careful note that no claim is being made here about what is "physically possible", since the existence of an external world is still bracketed by the Husserlian *epoché*. As long as a result fits into the context of the given basis, which includes my sense of personal identity, then we *must* include it when attempting to compute a probability distribution for the perceptual field. So long as both the ball under the table and the ball on the table are mathematically *possible* percepts to have, then we must include them both.

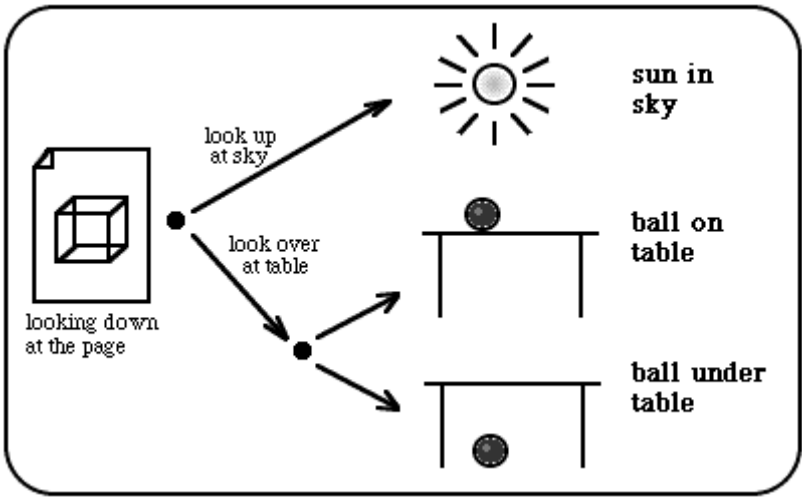

Fig. 9: Looking up from the page: 2 possible actions, 3 possible results.

We see immediately that a completely phenomenological theory of consciousness, while it may yet find a place for the material "world", is surely never going to be a theory of One True World. Already, we see that it can only be a theory of many possible worlds. This conclusion is not *quite* justified, even given (*7), for it is always possible that were we to fully formalise consciousness, we would find that there actually is only one possible result of looking at the table. This seems hardly credible, however. Surely, although we may see the ball *on* the table, it was still a mathematically possible perception to see it *under* the table. In any case, we will assume this viewpoint for now.

> **(*8)** Working assumption: it seems reasonable to assume that for most ordinary observations of what one normally calls the "world", there is more than one possible perception that could result.

What goes in the ket is, as we have seen, the thing-in-itself that our perception is *of*. In the case of "2+2=4", this thing in itself was something we could actually in practise detach from the unclear fringy field (something we could actually in practise know). But for most everyday perceptions like the ball, we see that such detachment is probably not possible to achieve in practise and perhaps not possible at all. For some percepts, it may not be possible to completely describe them without describing ourselves. If this is so, then the thing-in-itself may have to include the entire perceptual field 1-3, which as itself something that is unclear and dependent on self-identity may be even larger than our immediate conscious field 1-2. As a completed formal system, it will thus contain our entire conscious experience as merely a part. Our practical inability to perform the detachment for most ordinary observations makes this a strong possibility. If this is the case for the ball-percept, then the thing-in-itself that |ball-percept> represents may be more aptly described as |the-entire-situation>, since no piece of it can be detached from any other (thus, it is not something we can in practise really know, since it is

bigger than us and contains us—although the super-beings may be able to know it).

We will call such entire-situations (containing entire perceptual fields within them) "worlds". We call it an "entire situation" or "world" instead of an "entire perceptual field" because this is the ket, the thing-in-itself that we want to make clear, so if not detachable from our entire perceptual field, it must be bigger than it, merely containing it:

world(look-ball(ball-on-table))	*(functional notation)*
< ball-on-table | look-ball | world >	*(bra-ket notation)*

Does this notion of "world" correspond to our common sense idea of an external world? To some extent, at least, it does. Note that the thing-in-itself is both bigger than us *and* in principle not separable from us. Thus its formal structure will determine the distribution of percepts in our perceptual field. If all our percepts, even unclear ones like the ball-percept, were completely contained within us like two-plus-two is, there would be no surprises when we applied a control operator like "look-ball". Since the ball-percept would be completely contained in consciousness, just like two-plus-two, we could perhaps work out for ourselves what the possible results of look-ball would be. But this is not the way it is. Sometimes, we expect the ball to be there and yet we see something completely different. The distribution of possible results, if fully formalized, would be a distribution of things-in-themselves that are bigger than us, so we cannot expect the precise distribution to be obvious to the intuition. In short, the world is too big and too complex to fully predict, even though its objects are completely defined in terms of our own self-identity-based perceptual field.

So, although it is too soon to jump to conclusions, we will go ahead and use the term "world" for this self-identity-based ket. And since the common sense material world is still bra-ketted (if the reader will excuse the pun), in accord with the *epoché*, we absolutely cannot go from our ket-based notion of "world" to the traditional notion of a material world that goes beyond what is perceivable.

Recall that the ball-percept *as a world* is the thing-in-itself, so if formalized to full clarity, it may well be extremely complex compared to the seeming simplicity of the unclear, vague ball-percept as it appears apodictically before consciousness. The earlier bra-ket expression formalized the entire conscious experience of observing a ball on the table, given that we turn our heads and look to see its position. To fully formalize the expression is to clearly know the percept (we may need to solicit the help of our superbeing friends). If we express it in the λ-calculus we achieve a high degree of distinctness. Nonetheless, we can achieve greater and greater adequacy of expression, or distinctness, by expressing the percept in as many different ways as possible, since taking whatever is common between all these things will give us the distinct "thing-in-itself" (at least in the unreachable limit of all possible languages). Thus the thing-in-itself can be seen as a kind of combination over all possible representations of it (in all

possible languages), as discussed in previous sections. By "combination" I simply mean whatever is meant by "take what is common amongst". Although doing this is essential to human cognition and our understanding of forms, we do not yet know *what* this operation is, since we cannot yet formalize our entire perceptual field. A complete formalized theory of cognition, possibly based on hierarchical perceptual control theory, might make this clearer. For now, we will just call it the COMMON() or INTERSECTION() or ∩() function:

< ball-on-table | look-ball | world >
< ball-on-table | world > = ∩(x) < ball-on-table | x | world >
where x here ranges over all possible ways of viewing the ball as on the table

We can drop the bra from both sides of the equation to get a more general expression of the world, or the entire situation:

Quantum Rule #2:
|ψ> = ∩(x) |x|ψ>

This expression says that the world, ψ, as a thing-in-itself is that which is common to all the different ways we could possibly attempt to control our perceptual field (where formal analysis is considered one kind of control).

Presuming that our entire-perceptual-situation, or world, is indeed a thing-in-itself that has objective form or truth and can be analyzed, then the action of some operators to achieve control may be more likely to produce certain resulting perceptual states than others. |world> will have a kind of probability distribution, depending on what operator we choose to apply. | world > should thus be quantifiable in terms of a particular probability distribution for some class of results.

By saying that there is a particular probability distribution, I mean that not all results are equally probable. A probability distribution requires more than just a collection of possibilities. A *possibility* field does not necessarily yield a nontrivial *probability* field. If all we have is a collection of various possibilities, there is no reason to suppose that any are more likely than any others. Nontrivial probability distributions require that the possibilities be grouped into "equivalence classes", sets of results that are considered, for the purposes of probability calculations, to be the same.

It is easy to see that there are natural-seeming equivalence classes for perception. If we recognize commonalities, or formal structure, in our perception, the basis language for our perception can form the basis of a probability distribution. For instance, in the space of Necker cube perceptions, we would group more than one possible figure as roughly the same result. If we are controlling for a particular point of view, trying to flip from one to the other, there are many possible results that we would accept as equivalent. Sometimes, this is a result of

conscious decision. We decide that certain results are similar enough to deserve to be classified together. But other times, it is simply because *different percepts can yield the same perception*. For instance, although there may be noticeably different percepts that we would treat as equally belonging to the "ball-on-table" class, there may also be many different, distinct percepts that *cannot* be distinguished, from our personal point of view within the core of consciousness. So for us, these are truly identical results (on the bra side), but in reality, they must be treated as different things-in-themselves (on the ket side). This appears to be the case, since much of the specific probabilistic structure of the world seems complex and arbitrary to the untrained eye. Although extensive phenomenological experiments with body-manipulations[17] have revealed numerous "laws of nature", no complete and clear rationale in terms of formal structures has ever been offered for these laws. They appear ultimately arbitrary. Although such laws might eventually be completely formalised, the point here is that this would take much work, and the result would hardly be intuitively obvious to the untrained eye. The fact that body-percepts have this unclear partly unpredictable, partly predictable structure, is what leads us to label the world "material" and presume that there is something about it that goes beyond what can be clearly and distinctly perceived. Nonetheless, I cannot completely justify this move, so I will mark it for later reference:

> **(*9)** Assumption of complexity: assume for now that in most ordinary observations, there are multiple distinct percepts that are indistinguishable to consciousness, and that the number of such distinct percepts is different for different perceptions. By indistinguishable, we mean that different percepts produce identical consciousnesses (*not* that any individual consciousness has had an opportunity to compare the percepts and find them identical).

We now have all we need to begin a recursive-function-based analysis of perception in terms of probability theory. The probability of the ball's being on the table given that we apply the look-ball operator, is equal to the number of ball-on-table possibilities that are in our personal-identity perceptual possibility field, divided by the total number of possible results of the application of the operator (including ball-under-table results, for instance). We will represent the number of ball-on-table results with vertical lines on either side of the expression, calling it the"magnitude" of the perception, as follows:

> **| <ball-on-table | look-ball | world> | =** *# of possible ball-on-table percepts, given that we look for the ball, and given our entire perceptual context, or world.*

---

[17]I refer here to scientific experiments, which have established laws such as those of Newton, Einstein and Heisenberg. It is a mistake to see these as empirical and therefore outside the realm of phenomenology. Such experiments are fully phenomenological; they simply rely more heavily on body-manipulations than imagination-manipulations. Their main practical difference from thought (i.e., imagination) experiments is that they are very difficult to set up and carry out, and so we must accept the word of others as to their outcome (which would be another unjustified move and require an asterisk (*11)). However, there is no reason in principle we could not carry them out for ourselves and reproduce the results, just as we reproduce our own imaginary results.

So the probability of seeing the ball on the table when we look is:

$$\mathbf{P(ball\text{-}on\text{-}table \mid look\text{-}ball \mid world)} = \frac{\mathbf{|\langle ball\text{-}on\text{-}table \mid look\text{-}ball \mid world\rangle|}}{\mathbf{|\,|look\text{-}ball \mid world\rangle|}}$$

In general:

Quantum Rule #3:

$$\mathbf{P(x|\psi)} = \frac{\mathbf{|\langle x|\psi\rangle|}}{\mathbf{|\,|\psi\rangle|}}$$

We know from probability theory (a recursive theory) that the sum over all probabilities in a distribution is 1 (which is easily verified in imagination):

Quantum Rule #4:
$\mathbf{\Sigma_i\, P(x_i|\psi) = 1}$

Another completely intuitive result we can take from probability theory is:

Quantum Rule #5:

$$\mathbf{P(x|\psi) = P(\psi|x)}\ \frac{\mathbf{P(x)}}{\mathbf{P(\psi)}}$$

Of course, the two sides of rule #5 might not be as equally likely to be tested by experience. The chances that we could be checking for the existence of a ball-percept given the existence of the overall context (or world), is much more likely than that we could be checking that the world exists, given that we've seen the ball.

We will use figure 9 as an example, and assume it exhausts all the possible perceptions that could result from applying the look-ball operator (which is absurd, but it simplifies our analysis). We will assume that figure 9 shows percepts as they are distinguished from one another by consciousness, and that in fact, while sun-in-sky is the only possible thing-in-itself that results from look-sky, there are actually four possible distinct results from look-ball:

$\mathbf{|p_1\rangle = \langle ball\text{-}on\text{-}table|}$
$\mathbf{|p_2\rangle = \langle ball\text{-}on\text{-}table|}$
$\mathbf{|p_3\rangle = \langle ball\text{-}on\text{-}table|}$
$\mathbf{|p_4\rangle = \langle ball\text{-}under\text{-}table|}$

Note that in going from ket to bra, the subscripted indices come off. Since the first three percepts are indistinguishable to consciousness, they represent the same perceptual field. They are equivalent as bras, yet completely distinct as kets. So given |world>, we apply "look-ball", and we experience one of four possible percepts. We cannot, of course, experience all four, since they represent two different personal identities, all of whom can remember being the single identity that made the observation. So while I might say naively that there was a 50-50% chance of either <ball-on-table| or <ball-under-table|, I would be wrong. The actual probabilities are:

P(ball-on-table|world) = 0.75
P(ball-under-table|world) = 0.25

This is completely different than what we might have naively expected. However, some of this probabilistic structure should be discoverable to us. A single observation will never tell us anything, of course, since we can only be one of the two resulting persons. But if we repeat the experiment many times, under similar enough conditions, we can observe the probabilities and use them to build a model of the percepts as things-in-themselves that will give similar results. Such is the nature of scientific experiment. Indeed experiments in quantum physics have shown the world to have this sort of probabilistic structure. The bra-ket notation was invented by P.A.M. Dirac as a formalism for quantum mechanics. Experiment has allowed an incredible degree of formalization of |world>, called the wave function of the universe, but the task is by no means complete. Of all the various formalisms for quantum mechanics (including Heisenberg's matrix formulation and Schrödinger's differential equations), only the Dirac notation allows one to proceed from perception to atomic physics and not the other way around. Only the Dirac notation is inherently phenomenological (although Dirac did not develop it in precisely these terms).

We thus conclude that the world, if it is a rational (*7), complex (*9) structure, cannot be deterministic with respect to our observations. There will always (or at least often) be more than one possible outcome of an observation, and thus more than one perceptual consciousness with an identity that remembers being the single "me" that chose to make the observation in the first place. Our consciousnesses therefore continually splits with every observation we make (or at least those that meet our complexity requirements). Since the world is defined in terms of personal identity, this means that the world itself is continually splitting as we make observations. This result is predicted by quantum mechanics, but is usually seen as a bizarre, inexplicable result. Only when quantum mechanics is developed as a phenomenology does the result seem natural and just what one would rationally expect.

However, one never sees the reverse, a consciousness splitting backwards in time—such as two consciousnesses merging into one. While this seems counter to the nature of consciousness, there is no reason to rule it out completely. So we simply note that it has not yet been observed,

whilst keeping our eyes open for counter-examples.

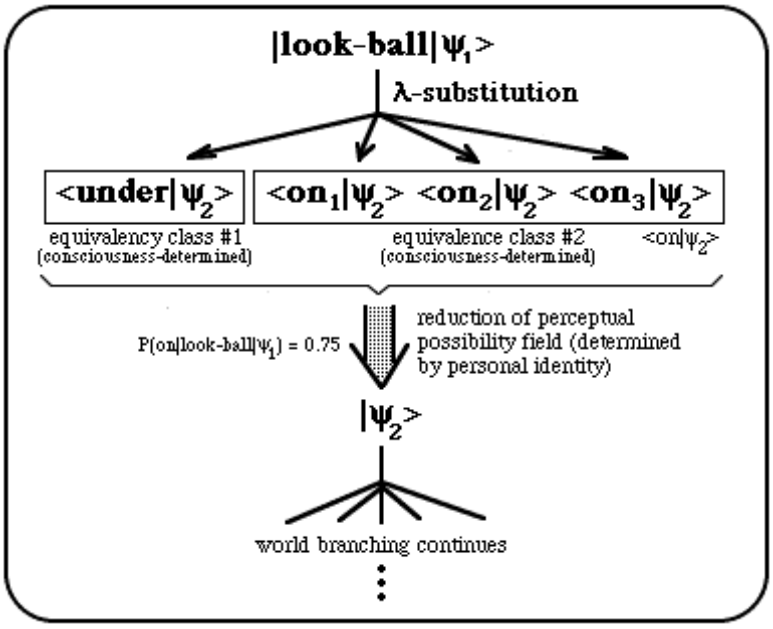

Fig. 10: Quantum Reduction of the Perceptual Possibility Field

Figure 10 shows the world-splitting that results in our example. Note that the world function, ψ, since it is identified with personal consciousness, must necessarily produce, as it runs on, all possible perceptions that the conscious being could have. ψ is called the "Schršdinger wavefunction" in quantum mechanics, and is what we are calling the perceptual possibility field. Given a complex world, with a nontrivial probability structure, the history of a conscious personal identity cannot be described as the running of the world, i.e., as a recursive structure, since that would be an enumeration of all possible histories, whereas any one person is only ever conscious of one of these histories. Since each time the world splits, more than one consciousness results, we only experience one of many possible futures, resulting in a subjective "collapse" or "reduction" of the possibility field. In quantum physics, this reduction is called the "collapse of the wavefunction", and it is a matter of great controversy as to whether the other worlds are truly "real" or if somehow there is something more to the world than the formalism of quantum mechanics. Treating quantum theory as a phenomenology gives a much different picture than the traditional formulations, heavily favouring what is known as the "many-worlds" or Everett interpretation of the theory.[18]

But then why is it, you might be asking, that when I look at a table, and then look away and look back a few seconds later, the table is still there? Why do I only experience a world in which tables are stable and don't just disappear at random or turn into giant pink bunny-rabbits? If these

[18]Hugh Everett, III. "'Relative state' formulation of quantum mechanics," In: *The Many Worlds Interpretation of Quantum Mechanics,* DeWitt & Graham (Eds.), pp. 141Ð149. Princeton U. Press, Princeton, 1957, 1973.

"weird" events are possible conscious states for me to have, then why don't I ever experience them? The answer quantum physics gives us is that these possibilities *are* in fact really there in your perceptual possibility field 1-3, even the giant pink bunny rabbit. But for large, complex objects like tables, such "strange" possibilities have very low probability, while what we think of as stable possibilities have much higher probability. Only when we are dealing with simple objects with little or no inner complexity, such as electrons and photons, do the probability distributions become uncomplicated enough to produce consciously noticeable effects that are "weird" and nondeterministic. So even the table's turning into a pink bunny-rabbit is allowed in quantum theory. But tables turn into rabbits in only a vanishingly small percentage of worlds, so although it happens in *some* worlds, we never observe it in practise.

Conclusion

Although we have had to make assumptions along the way, we have derived from basic phenomenological principles some of the fundamental concepts of both feedback control theory (the backbone of all biology) and quantum mechanics (the backbone of all physics). Let it never be said that Husserl's dream of founding all the sciences on phenomenology[2] was unrealistic. It is clear that both biology and physics can benefit from a reformulation in phenomenological terms. Before Powers[6] and others developed perceptual control theory, there was a non-percept-based version of feedback control theory that was far less phenomenological. Likewise, the highly phenomenological character of quantum physics so unnerved a scientific community unprepared for it that most in the field still believe today that there must be something fishy about the whole thing.

Whether you agree with the specific way I have developed these theories in this paper or not, I think it is clear that developing a science from the ground up as a rigorous phenomenology yields a completely different set of standards by which theories are judged as elegant and/or explanatory. From the phenomenological view, the explanatory power of a theory must ultimately lie in its ability to explain our immediate conscious experience, to satisfy our sense of wonder. Extrapolation to an external world is only permitted if there is a real reason for it based on experience. Rather than eliminating the world, or reducing it to something mental, we have put it in its proper place, and have given it powers beyond what is apodictically evident *only* when such powers seemed necessary in order to explain that very apodicticity.

As such, this view of the world is both the most extreme kind of rationalism, refusing to accept anything that cannot be clearly and (relatively) distinctly conceived of, *and* at the same time the most extreme kind of empiricism, as it grounds all that is rational in terms of experience, and what it views as requiring rational explanation is restricted to what is experienced. Radical rationalism and radical empiricism meet in the phenomenological method.

Not that it is impossible for there to be more to the world than what is justified by the

phenomena. But if there *are* such unknowable what-nots, we shall never know them, so how can they have any place in our philosophy? Even if, in the future, our experiments tell us there must be something "physical" about the world that is nonrecursive, this would only be a refutation of assumption (*7), not a refutation of the more general phenomenological method, which does not require (*7) at all. The core of the phenomenological method is its insistence on eradicating prejudice from science and philosophy. We cannot and will not accept the existence of nonrecursive entities until experience gives us reason to do so. But likewise, we shall not take our lack of evidence to imply the nonexistence of such entities, either. We simply take no stand on the existence of such mysterious substances. We refuse to assent to either their existence or their nonexistence. To paraphrase Wittgenstein, that which is inaccessible to consciousness must be passed over in silence.