Undecidable Long-Term Behavior in Classical Physics: Foundations, Results, and Interpretation.

Matthew W. Parker

Advisor: Howard Stein, Professor Emeritus of Philosophy

ABSTRACT

The behavior of some systems is non-computable in a precise new sense. One infamous problem is that of the stability of the solar system: Given the initial positions and velocities of several mutually gravitating bodies, will any eventually collide or be thrown off to infinity? Many have made vague suggestions that this and similar problems are undecidable: no finite procedure can reliably determine whether a given configuration will eventually prove unstable. But taken in the most natural way, this is trivial. The state of a system corresponds to a point in a continuous space, and virtually no set of points in space is strictly decidable. A new, more pragmatic concept is therefore introduced: a set is decidable up to measure zero (d.m.z.) if there is a procedure to decide whether a point is in that set and it only fails on some points that form a set of zero volume. This is motivated by the intuitive correspondence between volume and probability: we can ignore a zero-volume set of states because the state of an arbitrary system almost certainly will not fall in that set. D.m.z. is also closer to the intuition of decidability than other notions in the literature, which are either less strict or apply only to special sets, like closed sets. Certain complicated sets are not d.m.z., most remarkably including the set of known stable orbits for planetary systems (the KAM tori). This suggests that the stability problem is indeed undecidable in the precise sense of d.m.z.

Carefully extending decidability concepts from idealized models to actual systems, we see that even deterministic aspects of physical behavior can be undecidable in a clear and significant sense.

ACKNOWLEDGEMENTS

Many people facilitated this project in important ways. I would first like to honor my tireless, generous, encouraging, brilliant, and absurdly modest dissertation committee: Howard Stein, William Wimsatt, Robert Batterman, and Wayne Myrvold. Each member deserves special thanks: Professors Batterman and Myrvold for voluntarily serving on a committee outside their own institutions, Professor Wimsatt for handling all of the business details, and most of all Professor Stein for providing very careful and thorough critiques during his retirement and in the face of physical limitations and health concerns. David Malament was my advisor earlier on and has continued to provide support in various ways.

One year of research and writing for this project was funded by a generous fellowship form the Mrs. Giles Whiting Foundation. I must again thank my committee for their enthusiastic efforts to help me secure that fellowship. Another year was supported by a Shaw Fellowship, and a summer's preliminary research was supported by a Mellon Pre-Dissertation Fellowship. One year was generously funded by H. Waid Parker, my proud and occasionally nagging father, with the unknowing help of Aunt Olga Waid.

I would like to thank the University of Western Ontario and École Normale Supérieure de Lyon for their hospitality at conferences that contributed immeasurably to

my research. Western Ontario introduced me to Wayne Myrvold, a major catalyst in the development of this work. ENS Lyon introduced me to several mathematicians working at the frontier of computable analysis and is due special thanks for unexpectedly obtaining a travel grant for me.

Gaspare Genna graciously hosted me for two work retreats, at which I did some of my first original mathematics. Douglas Herman hosted another. Jim Guszcza and Shantanu Dutta have provided warm support and advice. The late Donald Kalish never knew of this project, but his outrageously kind and diligent teaching, his political activism, and his lust for life were great inspirations to me. It is because of him that I love logic.

Most of all, I must thank my wife, Laurel Parker. It was she who helped me to find the inspiration and ambition to enter philosophy, much to her detriment. I and my dissertation have been tremendous inconveniences to her, often making daily life a trial and sometimes obstructing major life plans. Nonetheless she has always shown startling enthusiasm and effectiveness in lifting my spirits when the road was rocky, and now, as the project reaches completion and my professional career begins, her pride and excitement are my greatest rewards.

I dedicate this dissertation to her.

MOORE'S WISH

At the end of his ingenious 1991 paper, Cristopher Moore writes,

> It is this author's wish that no one derive, directly or indirectly, military benefit from this work. Please copy this wish if you cite this paper.

I am happy to oblige, and I share his wish wholeheartedly.

TABLE OF CONTENTS

LIST OF ILLUSTRATIONS

Figure

LIST OF ABBREVIATIONS

cl	closure
d.i.b.	decidable ignoring boundaries
d.m.z.	decidable up to measure zero, or decidable "mod zero"
dom	domain
ext	exterior
int	interior
i.t.	indefinite-term
KAM	Kolmogorov, Arnol'd, and Moser
NET	neighborhood enumeration theorem
Pr	probability
r.a.	recursively approximable
ran	range
r.e.	recursively enumerable
r.m.	recursively measurable
TM	Turing machine
TTE	Type-2 Theory of Effectivity
TUP	topological use principle
λ	Lebesgue measure

INTRODUCTION

> It is unphilosophical to suppose that, to any given question (which has any clear meaning), investigation would not bring forth a solution of it, if it were carried far enough.
>
> Charles Sanders Peirce, "How to Make Our Ideas Clear"

0.1. The situation

Several authors have suggested that the very long-term behavior of certain physical systems, or more precisely, theoretical *models* of physical systems, cannot be predicted by any systematic calculation (Moser 1978, 67-68; Wolfram 1985; 2002, 755, 1138; Moore 1990, 1991; Pitowsky 1996; Sommerer and Ott 1996). These models are *deterministic and classical*, i.e., non-quantum. Their future states are entirely determined by their past states, and their supposed unpredictability is not due to any ontological indeterminacy but only to the lack, and the supposed impossibility, of a systematic means of calculation. Further, this supposed unpredictability is different from what is popularly called chaos. Chaos means that we need *extremely accurate* knowledge of present conditions in order to make even rough predictions of the finite future. The claim under consideration here is that even *perfect* knowledge of the present would not enable us to predict the *infinite* future—whether, for example, a planet will *ever* escape the solar system. Inaccurate initial data are not to blame; there is just no adequate method of calculation, and none is even possible.

Such claims demand clarification, for taken in the most natural way they are trivial. The states of the systems in question are represented by real numbers, or points in a continuous space, and the non-computability claims concerning these systems reduce claims that for some set of points that interests us, there is no method to determine whether a given point is in that set. They are claims that some set of points is *undecidable*. However, the most obvious notion of decidability for set of real numbers, or of points in a continuous space, turns out to be almost unsatisfiable. As we will see, only the most trivial sets—the empty set and the entire background space—are decidable in the obvious sense. Most authors writing on decision problems in physics seem to have overlooked this fact,[1] and as a result, it is not immediately clear what sort of non-trivial undecidability, if any, the systems they discuss might suffer, nor what significance such undecidability may have.

This raises a number of general questions. Are there inherently unsolvable problems built into even our simplest deterministic theories of the world? Is the behavior of some models undecidable in a meaningful sense? What after all *is* a decidable set of real numbers? Is the idealized problem of the stability of the solar system undecidable, as some have suggested (e.g., Moser 1978; Wolfram 1985, 2002)? Is undecidability only a property of models, or could there be actual systems in the real world exhibiting, in some sense, objectively undecidable behavior? Would this imply some form of non-

---

[1] Wolfram (2002, p. 1138) at least acknowledges the difficulty, and since he suspects that space and time are ultimately discrete, for him there may be no such difficulty.

determinism, at least for those who hold strongly empiricist or constructivist views, as Pitowsky (1996) has suggested? Or is it only our knowledge that is limited? Are there, in contradiction to remarks like Peirce's above, questions about the world, with clear meaning, to which no amount of investigation would ever bring forth a solution? These are some of the broader questions that motivate the present investigation.

0.2. What is a decidable set of real numbers? A pragmatic approach

You are free, therefore choose—that is to say, invent.

Jean-Paul Sartre, *Existentialism and Humanism*

There are stories about Karl Popper's brusque participation in seminars. Apparently, when someone would announce a title of the form, "What is *X*," Popper would immediately interrupt, " 'What' questions are completely wrong, misguided."[2] This attitude seems only a little more dismissive than that of the later Wittgenstein, for whom the philosophical "What is *X*?" could express little more than a mental discomfort, best answered by displaying the *grammar* or *use* of '*X*' (cf. Wittgenstein [1933-4] 1964, pp. 1, 26). The present investigation is somewhat sympathetic to such views. When philosophers approach a question like "What is a person?" or "What is art?" as a great mystery merely because they can imagine cases that defy classification, and when they seek the solution in some recondite essence, God-given demarcation criterion, or anything

[2] This is according to John Watkins, quoted in the popular book *Wittgenstein's Poker* (Edmonds and Eidinow 2001, 176).

other than human conventions, they seem to make words their masters rather than their tools.

However, we suppose here that merely identifying these pitfalls might not bring philosophy to an end. Perhaps having skirted such traps we can proceed with valuable philosophical work. After all, even distinctly philosophical "what is" questions are not without motivations. They may express a reasonable desire to extend established concepts, beliefs, and values to difficult cases.[3] To throw up our hands and deny that such questions have any meaning at all would again be to fall slave to words, by forgetting our power to *give* them meaning.[4] If an old concept does not fit a new case, one option is to invent a refinement or modification of the concept, one that can either solve practical problems or satisfy some curiosity more specific than the naïve "what is" question. Faced with a case where the question "Is *a* an *X*?" seems to have no uniquely

---

[3] As examples, consider the practical implications of concepts like *person* and *marriage* for current ethical and legal debates. Though it is a mistake to focus such debates on the meanings of words, debates over the words reflect a need to clarify or reinvent concepts and values.

[4] This remark is not aimed at Wittgenstein, for he writes, "[W]ords have those meanings that we give them" ([1933-4] 1964, 27), and rather than deny that a question "What is *X*" has any meaning, he discusses the various uses of '*X*,' and points out "family resemblances." However, he excludes conceptual engineering from philosophy. He tells us that philosophy "leaves everything as it is" (1953, 49-50), and that it cannot or need not improve on ordinary language (ibid.; [1933-4] 1964, 28). For Wittgenstein, philosophical work is finished when our *present* usage has been laid bare and confusions removed. Such a strict *laisez-faire* policy seems somewhat arbitrary. Philosophers are entrusted with the tasks of evaluating science and relating its technicalities to commonsense notions and traditionally philosophical issues like determinism and predictability. Such duties are just as much "heirs of the subject which used to be called 'philosophy'" (Wittgenstein [1933-4] 1964, 28) as Wittgenstein's own investigations. It would be impractical to separate these critical and interpretive duties from the constructive suggestions that might arise from them, or that might even contribute to them by showing that more useful concepts exist. Here we study and engineer precise mathematical concepts in hopes that they may shed light on philosophical questions after all.

correct answer, the philosopher might ask, "What would I really like to know about *a*? What motivates the question?" and define concepts accordingly. If these questions cannot be answered, *then* perhaps there really is no problem to discuss.

In mathematics, the sort of conceptual engineering proposed here is fairly common, as Lakatos illustrates ([1963] 1976). In an elaborate process of conjecture and proof analysis, concepts are often retooled to support tidy and powerful theorems. Summarizing and perhaps embellishing Lakatos's account, Peter Smith states several of the concerns that guide this investigation:

> Definitions in mathematics get shaped by a number of pressures. We may start with a cluster of informal basic results which we want our formal definitions broadly to sustain (some of the intuitive results may be non-negotiable; others may be up for possible revision). There is then, on the one hand, the desire for increasing generality, inclusiveness, abstractness. But on the other hand, we also want the defined concepts to feature in powerful theorems. Connectedly, we want there to be interesting relations to (refinements of) other, well-entrenched, mathematical concepts. (Smith 1998, 174)

It is worth noting that among the factors contributing to the value of a concept are the results that follow from it. The recovery of familiar results can show that a newly refined or generalized concept has not been modified too much, but also, unexpected results may attest to the importance of a modification. Even relatively simple results like those to be presented in Chapters 4 and 5 can speak to the merits of a concept, especially if, as in this case, they are applications to natural examples in other fields (here physics), and they illustrate the distinctions between different refinements of an old concept.

For an investigation such as this one into mathematics and physics, then, the present methodological preface might seem unnecessary. However, we are also concerned here with the notion of an effective procedure, a vague concept not far from the notion of rule following that preoccupied Wittgenstein (1953, 80 ff.). It is hoped that our technical—i.e., careful and detailed—investigations may shed some light on traditionally philosophical questions of epistemology and the limitations of reason.

Moreover, there has been a tendency even in the "exact" sciences to overlook the tearing and fraying that familiar concepts suffer when stretched beyond their usual applications, or to feel that one mended concept is the only natural extension of the original. The latter tendency is illustrated by the several nineteenth century mathematicians who, confronted with unexpectedly strange geometric solids insisted that those figures were not *really* polyhedra (Lakatos [1963] 1976, 14-21). "Essentialism," says Lakatos, "has been a permanent feature of definitional quarrels" (18), while such quarrels have themselves been papered over by the efficient but opaque Euclidean style of definition, theorem, and proof.[5]

The tendency to overlook conceptual fraying altogether is illustrated by physicists such as Wolfram (1985), Moore (1990, 1991), and Sommerer and Ott (1996), who have made claims of undecidability in classical physical systems without acknowledging that this is either ambiguous or trivial, due to the fact that the most obvious notion of a decidable set of real numbers is virtually unsatisfiable (see Sections 1.4.3 and 2.4). Others have noticed this fact and devised various relaxed notions of decidable or recursive set, as well as models of *computation* used to classify not only sets but real numbers, functions, and other structures in continuous settings. As some of their titles suggest, writings such as Blum et al.'s "Manifesto" (1996), Brattka's "The Emperor's

[5] We of course employ this style in some places (Chapter 3 and the appendix), but in light of and in service to the present pluralistic discussion of motivations.

New Recursiveness" (2003b), and others (Weihrauch 2000, Brattka 2003a) express distinct preferences for one concept or model over another.

Not every concept is precious, and some may not well serve our purposes, nor even the purposes for which they were proposed. One of our main goals will be to point out the special significance of one particular concept, namely that of *decidability up to measure zero* with respect to a given measure (called "decidability in μ" in Parker 2003). However, the arguments for of this particular concept and the critiques of other concepts presented here are valid only relative to particular purposes. Despite any apparent partisanship, they are intended to be taken in an ultimately pluralistic, constructive, and pragmatic spirit.

CHAPTER 1

ON THE INTEREST OF INDEFINITE-TERM BEHAVIOR IN PHYSICS AND OF UNDECIDABILITY THEREIN

1.1. Introduction

The claims of undecidability cited in the introduction concern the behavior of systems in the unbounded future. They concern properties involving quantification over all time, such as the property that at *some* time the state of a system will enter a certain set, or that the system will approach a given set of states asymptotically. Let us distinguish such properties from properties of behavior over some finite time, however long, by dubbing the former *indefinite-term* (i.t.) properties and the undecidability of such properties *i.t. undecidability*. There are also indications of *finite*-term non-computability in classical physics (e.g., Pour-El and Richards 1981), but here we focus on i.t. decision problems. One of the purposes of this chapter is to show that though i.t. undecidability is an abstract and theoretical matter—even more abstract than finite-term non-computability—it is worthy of a philosophical investigation.

There have been suggestions to the contrary. Philosopher Wayne Myrvold writes,

> Some of the predictions of a theory concern the value of measurable quantities. Others may concern the long-range behavior of the system, such as whether the system will ever leave a certain volume of phase space. We should, therefore, distinguish between predictions which are testable by experiments taking place in

> a pre-defined, bounded region of space-time, and predictions which are not. Only the former should be considered measurable predictions. This is significant because, if the dynamics of [a physical] theory permit the construction of a physical instantiation of a Turing machine, the corresponding halting problem will arise, so that certain features of the long-range [i.t.] behavior of the system will not be an effectively computable function of its initial state. Turing machines have been constructed (conceptually), not only out of electronic components, but of colliding billiard balls and also quantum systems [citing Fredkin and Toffoli 1982 and Feynman 1986]. Such non-computability is routine, and hardly counts as an instance of the physics outstripping effective mathematics, as, in such systems, the state of the system at any given time is an effectively computable function of the initial data. (1994, 1995)

This at least *seems* to suggest that i.t. predictions are empirically empty and i.t. undecidability is trivial. Some of Myrvold's opinions on this topic have changed, and anyway the real purpose of his remarks seems not to be to dismiss i.t. undecidability absolutely, but to limit the scope of those writings in which his remarks appear. We will not attack Myrvold as if he were committed to the claim that all i.t. undecidability is uninteresting. Nonetheless, he raises reasonable doubts that should be addressed.

In part, this chapter will serve to provide some historical background, on computability, the i.t. problem of the stability of the solar system, and suggestions that it and related problems are unsolvable. However, our main purpose here is to show that i.t. undecidability is of some interest. We will briefly consider the interest of highly abstract propositions in general, and of corresponding undecidability results, by analogy with Hilbert's philosophy of mathematics and the impact that undecidability results had on his research program. We will then see (1) what kinds of interest i.t. predictions related to the stability of the solar system may have, (2) that such predictions are testable in various ways, (3) how the appearance of *unsolvability* in problems related to the stability of the solar system has influenced the direction of research, (4) that i.t. undecidability is not as trivial as Myrvold's (and Stephen Wolfram's) comments seem to suggest, and (5) that there is at least some suggestion, due to Pitowsky (1996), that i.t. undecidability has metaphysical implications. We will not attempt to eliminate all doubts about the interest of undecidable i.t. behavior, but to disarm such doubts of their immediate force, for a full evaluation of the significance of i.t. undecidability can only be made after more rigorous definitions and results have been established. We will do so in Chapters 2 through 5 and make a more earnest attempt at interpretation in Chapter 6.

1.2. Abstraction, unbounded quantification, and undecidability

Indefinite-term propositions about mathematical models of physical systems are highly abstract idealizations. Though the significance of such propositions may not be immediate or concretely practical, it may be genuine and considerable nonetheless. At least for some individuals and groups, abstract propositions, especially if *suggested* by the concrete world, have inherent interest regardless of their practical implications. In the philosophy of science, to understand what we can and cannot predict, in exactly what sense, and why, are basic desiderata, and hence the study of non-computability in the i.t. behavior of abstract models seems important for the sake of thoroughness alone. Moreover, abstract idealizations can serve as heuristic guides or even direct shortcuts to more concrete results. *Undecidability* results for idealized problems, then, can serve to direct research programs away from the fruitless pursuit of such shortcuts.

One illustration of these general considerations can be found in Hilbert's view of the role of unbounded quantification[1] in mathematics, his program to justify it, and what became of that program. (For our purposes this is perhaps more than an allegory, since i.t. propositions about models of physical systems *are* mathematical statements with unbounded quantifiers.) Hilbert regarded unbounded existential propositions in arithmetic as "ideal elements," constructs artificially introduced in order to complete and simplify a theory (e.g., [1926] 1983, 195). Other examples of ideal elements include imaginary and complex numbers, the completion of Euclidean space with "points at infinity," and the algebraicobjects dubbed "ideals" by Dedekind. Hilbert thought that we could justify using unbounded existentials and applying classical logic to unbounded quantifiers if we could prove by "finitary" methods (a vague notion) that the results would be consistent. He thought this could be proved as a matter of concrete fact, without further appeal to unbounded existentials, by treating a theory as a formal system of symbols and simple rules (ibid., 199-200). So construed, the consistency problem could be regarded as a matter of elementary arithmetic, the symbols being equivalent to numbers and the rules to arithmetic operations.

Hilbert's hopes were of course dashed by an undecidability result—Gödel's

[1] Unbounded quantification is the use of statements such as "There exists a number n with property P," without specifying a bound on n, as in "There exits a number *n less than one million* with property P." The meaningfulness of unbounded quantification had been criticized by finitists and intuitionists.

second incompleteness theorem ([1931] 1986)[2]—but not quite sunk. That theorem showed that no consistent theory of arithmetic contains a proof of its own consistency. Hence whatever a "finitary" proof of the consistency of arithmetic might be, it could not be carried out within axiomatic arithmetic, contrary to Hilbert's expectations. Yet, rather than abandon Hilbert's program altogether, some went on to pursue broader notions of finitary proof or modified versions of the program (see Kreisel [1958] 1983; Kleene 1986, 139-141). Thus Gödel's undecidability results helped to direct research away from certain dead ends and toward more fruitful avenues.

Gödel's theorems are not quite analogous to the kinds of undecidability results that concern us here; they do not explicitly state that there is no algorithm to decide membership in some set. In closer connection to our topic, Turing demolished another element of Hilbert's formalist program by showing that there is no solution to Hilbert's *Entscheidungsproblem*, i.e., that there is no algorithm to decide the set of valid formulae of first-order logic (Turing [1936-37] 1965). This was, in fact, an i.t. undecidability result. Turing argued it by means of what we now call Turing machines (TMs), and one idealization inherent in that model of computation is that a computation may take an

---

[2] It is less often mentioned that Gödel's *first* incompleteness theorem undermined the very *motivations* for Hilbert's program. This point was the occasion for Gödel's informal announcement of the theorem at a Königsberg conference of 1930 (Gödel [1931a] 1986, 200-203). His point was that consistency is not justification enough for the application of an axiom or rule. "Contentual" considerations (such as the fact that a certain sentence is true if and only if it is not provable from the standard axioms of arithmetic) can establish arithmetic truths not proved or disproved by the standard axioms. Hence, adding the negation of such a truth to the axioms yields a theory that is consistent (if the axioms are consistent) but not true of arithmetic proper, i.e., the intended model. So even in mathematics, there is a real question of truth beyond that of consistency, on Gödel's view.

arbitrarily long time. The halting problem (to determine whether or not a given TM will *ever* yield an output) is thus an i.t. problem, and the unsolvability of the halting problem—*the* classic non-computability result—is an i.t. undecidability result.[3] Turing derived the unsolvability of the *Entscheidungsproblem* from this, and thus i.t. undecidability results directed research away from another problem that would never have been solved.

The point of this discussion is not to endorse or deride Hilbert's philosophy of mathematics. Rather it is to suggest a tentative picture of the usefulness of (1) propositions about the unbounded future of a system and (2) knowledge about the decidability of such propositions. It suggests that, just as Hilbert regarded unbounded existentials as shortcuts to finitary results, i.t. claims may derive their practical significance from their function as shortcuts to finite-term predictions. The practical value of an i.t. *undecidability* result, then, would lie in its power to direct research programs away from the pursuit of some such shortcuts (and perhaps toward more promising ones). However, this practical assessment is compatible with realist views that should not be dismissed out of hand: that claims about the unbounded future may represent genuine truths, either about abstract models or about the physical world, and

---

[3] Of course, this i.t. undecidability result concerns a system with discrete states rather than a real-valued system.

Conceived in terms of Turing machines (or in terms of any process of computation in which similar steps take equal time), decidability and undecidability themselves are i.t. properties: a property P is decidable if there is some TM that *eventually* determines whether a given whole number n has P, i.e., after *some* finite number of steps. However, not all undecidability results are i.t. undecidability results; i.t. undecidability is defined here as undecidability *of* an i.t. property.

that such truths may have genuine value aside from any practical applications. Hence, I suggest it is prima facie conceivable that i.t. undecidability results may point to some genuine and valuable truths that we cannot know. To determine whether this is in fact the case is just another reason to study the matter more carefully.

1.3. The interest and usefulness of indefinite-term predictions

1.3.1. The stability problem and Universal Gravitation

One long-standing i.t. problem, and one that has influenced much development in mathematics, is that of the stability of the solar system. Despite the appearance of fixed regular orbits, the endless tugging of every planet on every other raises a serious and difficult question as to whether a planet could ever be pulled out of its present orbit and flung out of the system or smashed into another. To approach this problem, we model the planets as point-like particles under Newton's theory of Universal Gravitation (UG). If none of the planets in the model ever escape or collide,[4] we say the system is *stable* (and in this dissertation, 'stable' should be taken in this sense unless otherwise indicated).

This problem should be distinguished from the finite-term n-body problem: to find the specific positions of n gravitating point masses at any *finite* time given their initial conditions. We have methods of numerical approximation that effectively compute

[4] The collision of point masses is inherently unlikely. The set of initial states of a system of n point masses leading to a collision has Lebesgue measure zero (Saari 1971-73), which is usually assumed to imply probability zero. The question of collision is part of the stability problem as it is usually framed. However, we, like several other researchers (e.g. Poincaré 1890;

the finite-term behavior of a system at least up to any point of collision or other singularity.[5] The indefinite-term stability problem, though, remains unsolved even in the non-singular case. We should also distinguish between the idealized *mathematical* problem of the stability of the solar system, which I will call the *idealized stability problem* or simply the *stability problem*, and the question of the actual fate of our actual solar system, which concerns bodies that are neither points nor spheres and involves factors other than gravitation. The latter I will call the *actual stability problem.*

One of the theses of this chapter is simply that some i.t. problems are interesting. The idealized stability problem is interesting in part because it resonates with broad cosmological concerns. Though it is an austere and ultimately unrealistic simplification, it is also symbolic of larger questions about the mutability of the heavens and the fate of mankind. To ask whether planets will escape or collide directly threatens the Platonic (*Laws* X, 898, 903) and Aristotelian (*De Caelo* 270ᵃ12-ᵇ24, *Metaphysics* 1071ᵇ3-1075ᵃ19, 1075ᵇ37-1076ᵃ4) vision of a perfect and cyclical heaven furnished by divine providence. Today there is little question of such perfection. We have long been confident that the actual solar system is ultimately unstable; that the faint resistance of interplanetary gasses and the friction of tidal motions will gradually drag the planets toward the sun (Poincaré 1898), and the sun will engulf several planets before finally dying out. Yet such

---

Arnol'd 1963; Wang 1991), will mainly be concerned with escape.

[5] The existence of such methods follows from the constructive existence and uniqueness proof for solutions of sufficiently regular differential equations (see Earman 1986, 117). An explicit power series solution to the finite-term problem (again, giving orbits up to singularity) is given by Wang (1991).

considerations might seem accidental. The idealized mathematical stability problem asks whether, even ignoring such details, catastrophe is inevitable. More than a metaphor for the question of our fate, it speaks to the *necessity* of the solar system's ultimate demise.

It is also at least loosely related to a historic disagreement between Newton and Leibniz. Newton suspected that under his own theory of gravitation, the actual solar system would eventually deviate from its present form. At the end of the *Opticks,* he wrote,

> [B]lind Fate could never make all the Planets move one and the same way in Orbs concentrick, some inconsiderable Irregularities excepted, which may have risen from the mutual Actions of Comets and Planets upon one another, and which will be apt to increase, till the System wants a Reformation. ([1704] 1979, 402)

Taken at face value, this passage seems to suggest that, due to mutual gravitation between the various bodies, the current configuration of the solar system is unstable in the sense that it will increasingly deviate from approximately circular, concentric, and coplanar orbits. We should be careful not to equate this with the sort of instability we have been discussing: escape or collision in an idealized system of point masses, due to gravitation alone. Newton does not raise the question of escape or collision here, and he may have had in mind causes of disruption to the system other than mutual gravitation. He frequently mentions the possibility of a gradual decay of motion due to the faint resistance of a very rare interplanetary medium. Nonetheless, the passage explicitly refers to "the Actions of Comets and Planets upon one another," presumably due to gravitation. Furthermore, eventual escape or collision seems a natural conclusion. Increasing irregularities known in Newton's time included an apparent expansion in the orbit of Saturn and shrinkage in those of Jupiter and the moon (Berry [1898] 1961, 203, 256). Hence it should have seemed that without a "Reformation," Saturn might eventually escape, the moon might collide with the earth, and Jupiter might collide either with the sun or with another planet.

Yet, rather than suppose that the system would ever deviate far from its present configuration, Newton seems to have preferred to postulate that God would intervene to keep the system in order. Leibniz objected to this, writing,

> [T]he machine of God's making is so imperfect, according to [Newton and his followers], that he is obliged to clean it now and then by an extraordinary concourse, and even to mend it, as a clockmaker mends his work, who must consequently be so much more the unskillful a workman as he is more often obliged to mend his work and to set it right. ([1715] 1989, 320-321)

From Leibniz's point of view, Newton's proposal seemed to contradict God's omnipotence.

This illustrates one way in which an i.t. problem could conceivably bear on the tenability of a theory. If one were committed to the view that an omnipotent God created the solar system and wished it to retain its form forever, then a prediction that the solar system would eventually deteriorate in some way might raise some theological concerns, as in fact it did for Leibniz. On the other hand, a very strong i.t. result predicting that, even taking resistance and other factors into account, the solar system would indeed retain its form forever, would have exonerated UG from Leibniz's theological criticism.

Theology aside, an i.t. prediction derived from UG could also have helped to confirm or disconfirm UG with regard to observation, at least in principle. A prediction of actual stability, for example, would imply that no body will ever escape, and escapes are relatively easy to observe; if a body recedes far enough and fast enough from the center of the system, then we can be certain (within reason) that it will escape, and given UG we can determine just what is far and fast enough to guarantee escape. So, if UG implied stability, then an observed escape would disconfirm UG, if not refute it, while the continuing absence of an escape would tend to confirm UG. In effect, a stability prediction implies infinitely many finite term predictions, predictions of the form, "No planet will exceed velocity v and distance r at time t," for fixed values of v and r but infinitely many values of it. Each of these finite term predictions can then be compared with observations.

Granted, this way of testing UG might not have been the most convenient. Very precise approximations of the positions of the planets at specific times already lent much confirmation to UG in Newton's time (though these predictions were not always exactly right). Nonetheless, an analytic stability result could have provided infinitely many finite-term tests in one fell swoop, without endless calculation.

1.3.2. Stability in particle accelerators

A perhaps more convincing illustration of the usefulness of i.t. predictions is provided by a problem very similar to that of the stability of the solar system, namely that of stability in particle accelerators. For example, at CERN, the European particle physics laboratory, sub-atomic particles are accelerated inside a narrow torus, something like a

600 meter hula hoop. The particles must be held in nearly circular orbits by a magnetic field in order not to collide with the wall of the chamber as they circle many millions of times, and according to Moser (1978), even more times than the earth has circled the sun. Hence, the magnetic field must be engineered so that the orbits of the particles are at least very nearly stable. Moser (ibid.) reports that the *finite*-term problem of designing such a nearly stable system was found to be practically impossible even with the use of computing machines, due to rapid growth in errors during the calculation.

The difficulty was overcome by solving an *indefinite*-term problem. The theory of Kolmogorov, Arnol'd, and Moser (KAM theory) demonstrates analytically the existence of many bounded orbits in a large class of mechanical systems. According to Moser (1978), it has shown that for certain accelerator designs, the great *majority* of orbits will avoid colliding with the wall, not only for millions of cycles, but forever (at least in an idealized model).

This is not only useful for the construction of accelerators, it also shows again how i.t. results can be used to test theories. Suppose our theory predicts that only a very small portion of particles should ever collide with the wall of a certain accelerator. If in fact a very large portion of particles collides with the wall in finite time, then something is wrong with our theory. This is assuming that nothing is wrong with the accelerator itself, etc., but one will always have to assume auxiliary hypotheses in order to test another hypothesis.

1.4. The testability of indefinite-term predictions

1.4.1. Falsifiability

How can i.t. predictions be used to test theories, or for anything else, if they are, as Myrvold says (op. cit.), not measurable?

It is true that i.t. predictions cannot always be *conclusively* tested by experiments that take place in a pre-defined, bounded region of space-time. For example, we might attempt to test a prediction that the actual planets will never pass beyond certain bounds merely by observing them. If our prediction is correct, then however long we watch, we can never rule out, based on observation alone, the possibility that some planet will escape at a later time.

However, one should not conclude that i.t. predictions have no empirical content. As indicated in Section 1.3.1, many i.t. predictions *are* testable in a distinct and fairly obvious sense. A stability claim, for example, asserts that no planet will escape, and an escape is relatively easy to observe. For any given system of bodies, we can establish clear sufficient conditions for escape; if at any time we observe one body receding from the system's center of gravity at a sufficiently high speed, then we can be sure that the stability claim is false. Hence a stability prediction is falsifiable in a fairly strong Popperian sense; we can imagine observations that would seem to force us to reject it.

Of course, whenever we seem to be witnessing an escape, we can always deny it at the cost of some auxiliary hypothesis. We might suppose that our optical equipment is not working the way we had thought, or that the planet is shrinking and has changed color (to account for red shift), etc. In calling a stability claim falsifiable, we have in mind what Lakatos calls "naïve methodological falsificationism," which takes into account the role of auxiliary hypotheses and ceteris paribus clauses (Popper [1934] 1968; Lakatos 1970). On this view, if we observe what seems, given certain auxiliary hypotheses, to be an escape or collision, then the conjunction of stability with those auxiliary hypotheses is refuted.

Admittedly, even this is a discredited "naïve" view. One might well prefer a more sophisticated methodology in which a hypothesis is only truly falsified when some other view proves more fruitful, or in which nothing is *ever* absolutely falsified (cf. Lakatos 1970). But this is really aside from the point. Whatever methodology of science may be the best, a claim of stability in a planetary system is about as falsifiable as hypotheses get. It can be refuted so straightforwardly that there could be little serious concern about auxiliary hypotheses or competing claims; the idea of dismissing an observed escape by rejecting the theory of optics, blaming the equipment, etc., is intuitively implausible. The claim here is not that some hypotheses can be absolutely falsified, only that a stability claim can be falsified *just as* strongly as any.

Again, we do want to recognize a difference between the testability of finite-term and indefinite-term claims. I.t. hypotheses cannot be reliably and conclusively tested within a prescribed finite time. A claim such as "Mars will recede beyond Pluto by the year 2025" can essentially be verified or falsified by that date. A stability claim, on the other hand, cannot be directly verified, only falsified, and we cannot say when. Other i.t.

claims, such as instability claims, can only be verified and not falsified.

Note, though, that this places i.t. predictions in no worse a position than many of the important theoretical generalizations that pervade science, and perhaps better than some. Newtonian universal gravitation (UG), for example, being a thesis about all bodies at all times (hence the 'U'), can never be verified, firstly because we can never observe all bodies, and secondly because we cannot observe a body for all time. Furthermore, UG is more difficult to falsify than a mere stability prediction, for it has a lot more wiggle room with regard to auxiliary hypotheses. Whenever it conflicts with observations, one can suppose that other forces or an unseen body are to blame, and on occasion this has turned out to be right! Again, though, auxiliary hypotheses are a side issue. The main point is that the *i.t. character* of a stability claim does not make it any less testable than other universal generalizations.

1.4.2. Testing asymptotic behavior

There are some indefinite-term propositions, quantified both existentially and universally over all time, that are neither falsifiable nor verifiable. Yet these too may be testable in a sense. A claim that an orbit $x(t)$ will converge asymptotically toward a given set A of physical states, for example, takes the form

$$(\forall \varepsilon > 0)(\exists T)(\forall t > T)\ d(x(t), A) < \varepsilon,$$

where d denotes a measure of distance within the space of possible states. Such a claim cannot be verified or refuted by direct observation, for if $x(t)$ seems to approach A in the short term, it may still diverge later, and if it seems to diverge from A, it may return later.

Convergence and divergence claims may nonetheless have theoretical interest as well as practical heuristic value, just as stability claims do. The divergence of orbits can be important for understanding the predictability of a system, as in cases of deterministic chaos. Further, if we know whether or not an orbit converges to a given set in the long run, we are in a better position to guess about its finite-term behavior, and conversely, finite-term behavior can suggest i.t. behavior. There is after all such a thing as *appearing* to converge asymptotically, and in some contexts, this appearance can constitute genuine and explicitly quantifiable evidence of convergence.

Consider the technique Sommerer and Ott (1996) use to graph the basins of attraction in a dynamical system.[6] For each initial state in a 760×760 grid of states, each corresponding to a pixel on a screen, they simulate an orbit until it comes within a distance of 10^{-5} from either of two attractors, with a speed transverse to the nearest attractor less than 10^{-6}. Then they color the pixel corresponding to the initial state either black or white, depending on which attractor the orbit approaches.

Sommerer and Ott argue that this procedure is accurate, based on their analytic results in Ott et al. 1994 for another, more artificial system with similar dynamics. Suppose the state of this more artificial system lies at distance ε from an attractor A. Ott and company argue that the probability that the state will *not* approach A asymptotically is

[6] The *basin of attraction* β of a set A of states is just the set β of initial states such that the orbit of a system beginning in β will asymptotically approach A. An *attractor*, in Sommerer and Ott's usage, is essentially just a set whose basin of attraction has non-zero volume.

$K\varepsilon^{\eta}$, where K, $\eta > 0$ are constant. Hence, even though there is always a positive probability that a given orbit will not converge toward A, this probability approaches zero for orbits close to A. Ott et al. also argue that such scaling of basins is universal for a certain class of systems, including that of Sommerer and Ott (1996).

Thus, in this theoretical context, the i.t. property of asymptotic convergence to an attractor can be probabilistically confirmed or disconfirmed by simulating an orbit for a finite time. Further, if Sommerer and Ott's system, or one sufficiently like it, does accurately model some physical system, then asymptotic convergence in that system can be confirmed or disconfirmed by a finite *observation*.

1.5. The stability problem, unsolvability, and the direction of research

There have been several suggestions either that the problem of the stability of the solar system or some related problem is unsolvable. The responses to such speculations illustrate how unsolvability, or supposed unsolvability, can direct research programs.

Laplace may have been the first to propose that the problem of stability—at least that of the actual solar system—is beyond us. Such stability, he wrote, "is disturbed by various causes that can be ascertained by careful analysis, but which are impossible to frame within a calculation" (1878-1912, 7: 121). Though Laplace himself had contributed one of several "proofs" of the stability of the solar system to emerge in the late eighteenth century, his proof, like those of Lagrange, Poisson, and others, was not a rigorous proof of indefinite-term stability, even for the idealized point-particle model, but a partial result that accounted for observed irregularities at particular levels of approximation (Berry [1898] 1961). Simply put, Laplace showed that to a first approximation, the mutual influence of the planets would cause their orbits to vary in long periodic cycles, but with no net change. In particular he accounted for the variations known to Newton in the orbits of Saturn and Jupiter, and thus his results suggested that such irregularities would correct themselves (op. cit., 314).

Yet, Laplace did not believe he had proven the absolute stability of the actual world system, nor, as the above remark shows, that this was possible. Though he believed the entire universe could in principle be described in every detail by a single formula, the human mind, he thought, "would always remain infinitely removed" from the intelligence required to discover and apply such a formula ([1814] 1951). This recognition seems to have benefited the direction of Laplace's research, for he cites such limitations as the motivation for studying probability theory (ibid.). It thus appears that the impossibility of certain knowledge in some areas encouraged Laplace's contributions to probability.

It is sometimes said that Poincaré proved the n-body problem unsolvable, as Diacu (1996) notes, and this is also said of the stability problem. This is partly due to Poincaré's contribution to a famous mathematical contest in the late 1880s. Weierstrass proposed as one of the contest problems to find a convergent series solution to the n-body problem, which he expected to yield a rigorous stability proof. Far from fulfilling

Weierstrass's hopes, Poincaré's winning treatise squashed them (1890).[7]

Yet Poincaré did not prove either the stability problem or the finite-term n-body problem unsolvable. In truth he discovered three interrelated reasons to believe that those problems must be extremely difficult. Firstly, he showed that certain known series used to predict the positions of the planets diverge in the general case (though he emphasized that this did not diminish their usefulness for finite-term approximation). Secondly, he showed that the three-body problem could not be solved by finding *first integrals*.[8] Finally, he discovered in the three-body problem an incredibly complex web of orbits now known as the *homoclinic tangle,* a standard example in chaos theory. About this he later wrote,

> One will be struck by the complexity of this figure which I do not seek even to trace. Nothing is more proper to give us an idea of the complication of the problem of three bodies and in general of all the problems of Dynamics where there is no uniform integral and where Bohlin's series are divergent. ([1892-9] 1957, v. 3: 389)

He then remarked that to solve the three-body problem, it would be necessary to devise methods entirely different from those known in his time (391).

---

[7] More accurately, the *corrected version* of the winning treatise did so. The original prize-winning treatise actually contained a fallacious stability result. The dramatic story of these events is told in Diacu and Holmes 1996, and with greater mathematical detail, in Barrow-Green 1997.

[8] A first integral is a constant of the motion—a function on phase space that is constant on any one orbit. By finding a sufficiently well-behaved first integral, one can transform an initial value problem into another with fewer variables and thus in effect reduce the phase space to one of fewer dimensions. Some problems, such as the two-body problem (including the two-body stability problem), could be completely solved using such techniques. However, Bruns had shown in 1887 that the three-body problem, with its 18-dimensional phase space, had only ten *algebraic* first integrals, and Poincaré strengthened that result to exclude all other "uniform"

Such results might have suggested that the three-body problem was unsolvable, but Poincaré himself tried to disabuse the public of that misconception:

> Do I say therefore that the problem is unsolvable? This word has no meaning; we have known since 1882 that the quadrature of the circle is impossible with a ruler and compass, and yet we know π with many more decimals than any graphic construction could give. All that we can say is that the three-body problem cannot be solved with the instruments to which we are presently disposed; those which it will be necessary to devise and to employ in order to obtain the solution must certainly be very different and of a much more complicated nature. (1891, 4)

Hence, Poincaré's response to the difficulties he encountered is not abject despair, but to suggest a change of direction. Indeed, in later papers, Poincaré attacked related problems but in a rather different setting: the study of closed geodesics on convex surfaces (see Barrow-Green 1991, 167-171), and in his last paper on the topic, he used quite different methods based in algebraic topology (ibid., 169). One might also surmise that the entire qualitative approach to dynamics for which he is often lauded, i.e., his interest in global, topological properties of systems,[9] was a response to the growing feeling that many differential equations could not be solved explicitly (ibid., 29).

Another contribution to the stability problem came from KAM theory, which, as mentioned above, demonstrated for a large class of mechanical systems the existence of

---

(real analytic) first integrals.

[9] It is worth noting that such global, topological properties of systems are in general i.t. properties, and even in his 1890 paper, Poincaré makes much use of i.t. results such as the existence of periodic, asymptotic, and "doubly asymptotic" orbits. For Poincaré, such results formed an important foundation for understanding, a "solid ground on which to support oneself to set out on new conquests" (1891). This illustrates yet another value of i.t. results.

many *quasi-periodic* orbits, simple orbits confined to tori in phase space.[10] For planetary systems, this showed that many states lie on stable orbits, provided all bodies but one are very small. In that case, the stable orbits established by KAM theory form the great majority.

It is in an effort to argue the value of such analytic i.t. results that Moser makes his apparent undecidability claim. "The stability of undamped systems *for all time,*" he writes, "can not in principle be decided by finite calculations and *lies therefore beyond the range of calculating machines*" (1978, 67-68, Moser's emphasis).[11] In response to this difficulty, Moser touts KAM's partial results. "[O]ne is led to a *new concept of stability* in which the restriction applies only to the *majority* of certain orbits...[This] weakened concept of stability is very meaningful and satisfactory for the physical applications" (1978, 67, Moser's emphasis). Thus Moser follows Laplace through the probabilistic turn, relinquishing the quest for exact prediction where it seems impossible and settling for a result of probable stability.

In all of the cases cited in this section, an appearance of unsolvability has led research in new and more fruitful directions. In these cases, the apparent unsolvability may or may not have been genuine; there were no rigorous unsolvability results.

---

[10] See Chapter 5 for a complete definition of 'quasi-periodic.'

[11] This may not be *intended* as a strict undecidability claim. Moser may only mean that the usual techniques for computing finite-time predictions will not in general answer stability questions. However, we will see in Chapter 5 how KAM theory suggests that the stability of the solar system may indeed be undecidable in a well-defined and significant sense.

Nonetheless, these examples serve to illustrate the valuable role that the recognition of unsolvability can play in the direction of research.

1.6. Triviality

> I do not think your observation was so trifling; in fact it was quite ingenious. To make this clear I shall show that it is false.
>
> Imre Lakatos, *Proofs and Refutations*

1.6.1. The triviality of naïve undecidability

Myrvold's comment that certain i.t. undecidability results are routine suggests that they are consequently uninteresting. Yet Wolfram (1985, 2002) claims that undecidable i.t. behavior is common and seems to feel that this makes it *more* interesting. This appears reasonable; if some feature of our world or our theories is common, that is all the more reason it should be well understood.

One might, then, take Myrvold's passage to suggest that undecidable i.t. behavior in physical systems, such as that derived from the unsolvability of the halting problem, is well understood or even trivial. In one sense, undecidability in real-valued systems *is* trivial, but it is partly for this very reason that it is not well understood. To see this, we must consider some basic recursive analysis.

The rational numbers can be effectively coded as natural numbers, so there is a straightforward concept of a recursive function taking rational numbers to rationals.

From this we can derive an almost equally natural notion of a computable *real*-valued function. A function f on the *real* numbers is *Grzegorczyk-computable* if, roughly speaking, there is a recursive function on the rationals that approximates $f(x)$ to any desired accuracy, *given a rational argument q sufficiently close to x* (Pour-El and Richards 1983, 542; cf. Grzegorczyk 1955, 1957) (A precise definition is given in Chapter 3, Definition 3.2.6.) As in the case of recursive functions on the natural numbers, several other well motivated concepts have proved equivalent to Grzegoczyk-computability. Hence this seems to capture an especially appealing notion of computable function, and it generalizes naturally to functions on the space $\mathbb{R}^n$ of real vectors, i.e., n-tuples of the form $(x_1, x_2, \ldots, x_n)$.

Yet, it does not suggest a useful notion of a decidable *set* of reals or real n-tuples. The obvious definition yields almost no decidable sets at all. Let us call a set $B \subseteq \mathbb{R}^n$ *naïvely decidable* if its characteristic function

$$\chi_B(x) = \begin{cases} 1 & \text{if } x \in B, \\ 0 & \text{otherwise} \end{cases}$$

is Grzegorczyk-computable. But all Grzegorczyk-computable functions are continuous. The only subsets of $\mathbb{R}^n$ with continuous characteristic functions are the null set and $\mathbb{R}^n$, so only these two most trivial sets are naïvely decidable. On $\mathbb{R}^n$, naïve decidability is just triviality, and naïve undecidability is trivial.

Myrvold may have had this in mind when he called i.t. undecidability results routine. Yet, his reference to arguments involving Turing machines (TMs) and the halting problem would be unnecessary to make that point. Such arguments seem to promise something more than the trivial naïve undecidability, some significant undecidability that has not yet been clearly specified. Hence the conclusions of such arguments are apparently not well understood; at least no one has clarified them in print.

1.6.2 The non-triviality of indefinite-term undecidability

Myrvold and Wolfram both allude to physical instantiations of *universal* Turing machines, those capable of computing *any* partial recursive function on any argument. Such instantiations are not as trivial as those authors seem to suggest. There is one great difficulty in constructing a universal Turing machine: it requires unlimited storage capacity. A desktop computer is not a universal machine in the same sense as an abstract universal TM. The former has only a finite amount of memory and disk space, and consequently, it is not capable of computing an *arbitrary* recursive function on an *arbitrary* input. It can only compute a vast but finite number of functions with finite domains. Therefore its behavior is decidable; its finite-term and i.t. behavior can be predicted by a more powerful computer, or even by a finite list pairing each of finitely many inputs with an output, or with a symbol indicating that the machine does not halt on that input.

The hypothetical computers cited by Myrvold and Wolfram are similarly finite and predictable. Let us take Fredkin and Toffoli's (1982) billiard machines, for example. Fredkin and Toffoli show that for any recursive function *f* on the natural numbers, and for any natural number *k*, one can construct a hypothetical system of flat reflectors and colliding hard spheres that would in theory compute $f(n)$ for all $n \leq k$. (See Figure 1.) However, there is no *one finite* Fredkin-Toffoli machine that computes $f(n)$ for *all n*, for a Fredkin-Toffoli machine cannot even read an arbitrarily large input. The input to a Fredkin-Toffoli machine consists of elastic balls of equal size, moving at the same fixed speed. These balls code information only in virtue of their presence or absence at certain specified points. An input 1011001, for example, consists of four balls in place of the four 1's in the string, and three empty spaces corresponding to the zeros, as shown in Figure 1. Hence a single machine would have to be infinitely large in order to accommodate arbitrarily large inputs consisting of arbitrarily many balls. Fredkin and

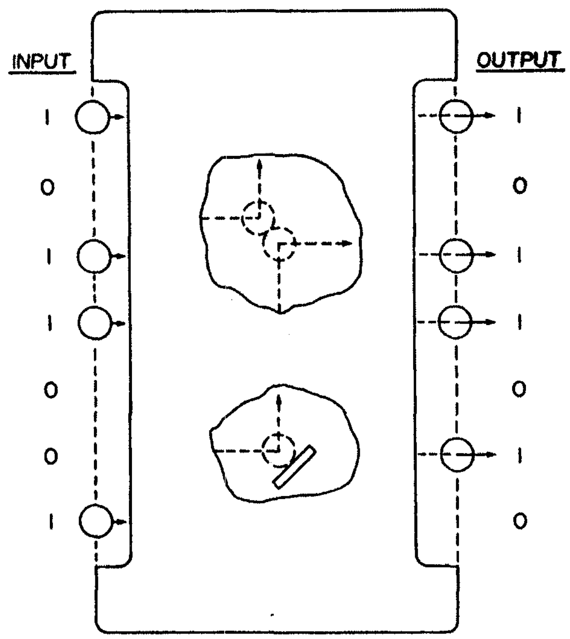

Figure 1.1. Fredkin and Toffoli's billiard computer (Bennett 1982). Computation is carried out by collisions between balls, directed by fixed flat mirrors. Input consists of balls entering the machine at any of finitely many specified points. In the example pictured, at most seven balls may enter. Hence there are only finitely many possible inputs, so no one such machine is equivalent to a universal Turing machine.

Toffoli's machines are universal in the sense that for any function on a finite range of inputs, *some* Fredkin-Toffoli machine will compute it. But no finite Fredkin-Toffoli machine is universal in the same sense as an abstract universal TM, only *one* of which suffices to compute *all* partial recursive functions on *all* integer inputs. Nor are Feynman's (1986) quantum computers universal in this sense; each bit of input to such a computer is "written" on a particle as a state $|0\rangle$ or $|1\rangle$. In order to be truly universal, such a machine would have to contain infinitely many particles.

Therefore the usual argument from the halting problem that there exists a *particular* universal TM with undecidable behavior does not apply to Fredkin and Toffoli's billiard machines, nor to Feynman's quantum computers—at least not if these machines are to have finite mass or occupy a bounded region of space.

Cristopher Moore (1990, 1991) has proposed a different kind of billiard machine that seems to circumvent spatial limitations. We will not go into great detail, but in Moore's machines, a single moving particle represents a bi-infinite string of symbols by virtue of the exact point at which it passes through a designated plane (in a direction orthogonal to the plane). The motion of this particle is directed by reflectors in such a way as to simulate a Turing machine (Figure 1.2). (This already seems ingenious, not trivial.) Given any TM, Moore shows, in general terms, how to construct such a billiard machine to simulate it. Hence some such machine simulates a *universal* TM, and by the unsolvability of the halting problem, its i.t. behavior is undecidable.

However, Moore's claim is not completely clear. Like other authors, he neglects

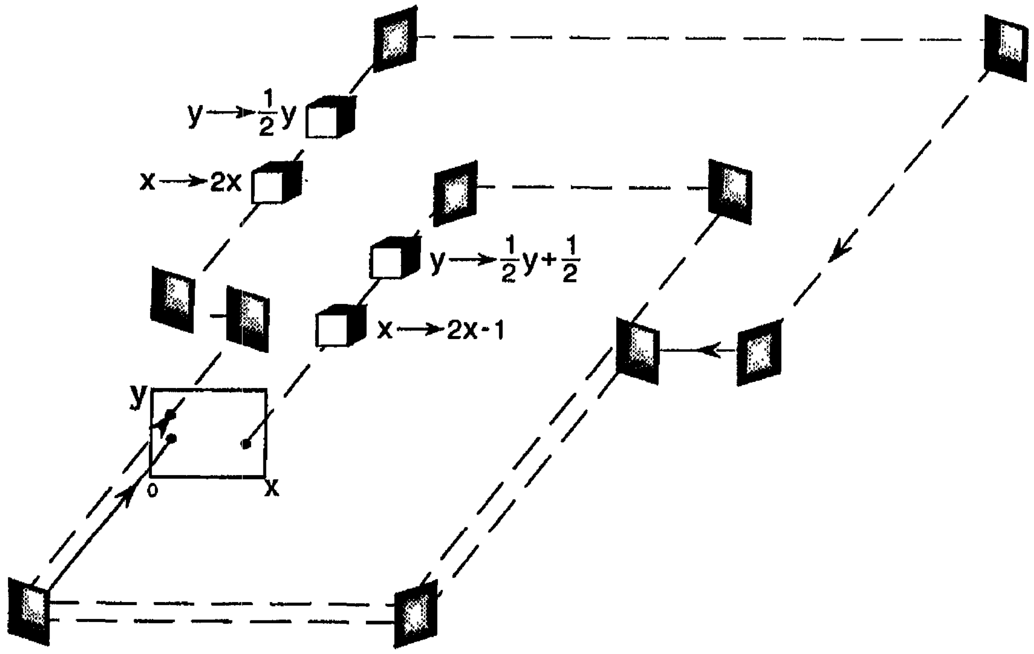

Figure 1.2. Schematic of a Moore-type billiard machine (Pitowsky 1996). The configuration of a Turing machine is coded in the x- and y-coordinates of a particle passing through the square shown. Small boxes labelled $x \to 2x$, etc., represent pairs of confocal parabolic mirrors, which shift the coordinates.

to define a non-trivial notion of undecidability for sets of reals. Also, Moore's argument appeals to a theorem that does not seem to apply in his context. Since the fundamental laws of classical physics, as well as currently accepted laws, are reversible, any candidate for a plausible model of a physical system should also be reversible. In particular, the function $\varphi(t, x)$ taking the state x of the system at time t_0 to the state at time $t_0 + t$ should be invertible, i.e., one-to-one. Moore attempts to ensure this by appealing to a theorem of Bennett (1973), which guarantees the existence of invertible, universal TMs. Given such a TM, Moore can construct a corresponding invertible billiard machine. However, *Bennett's theorem does not apply* to the kinds TMs on which Moore bases his constructions, namely *single-tape* TMs allowing arbitrary *bi-infinite* inputs. It applies only to TMs supplied with an infinite amount of *blank* tape. It would not help Moore to assume that his inputs are finite, or that they occupy only every second space on the initial tape. In order to be physically plausible, Moore's machines must be invertible for *all* initial conditions, not just those corresponding to certain sanctioned initial tape configurations.

There may well be remedies for this problem, but it is not *presently* clear that Moore's billiard machines can be made to exhibit any non-trivial sort of undecidability in real space. Moore also suggests other, more abstract forms of machines in higher dimensions, which might be physically more plausible (1991), but it is not *obvious* how to construct concrete instantiations of such machines, nor, again, that a significant kind of undecidability would result. Significant undecidability in a physical instantiation of a universal TM may indeed be possible, but it is not a trivial matter.

1.7. Indefinite-term undecidability and physical reality

Myrvold is right to point out that i.t. undecidability "hardly counts as an instance of the physics outstripping effective mathematics" in cases such as those in question, where indeed, "the state of the system at any given time is an effectively computable function of the initial data" (1994, 1995). However, if i.t. undecidability of a significant kind does exist, it may at least imply that a certain *ontology* in a sense outstrips our powers of computation, namely that ontology in which the world—past, present, and future—consists of a four- (or higher-) dimensional manifold onto which events are indelibly painted. This picture constitutes a kind of determinism, or perhaps better, *determinacy*, since it does not imply an ordered, law-like dependence among events, but only that there *are facts* about what events will occur in the future and when.[12] It is the tacit assumption behind Russell's insistence that the future "will be what it will be" ([1912-1913] 1963, 146), namely the assumption that there *is* something in particular that the future will be.[13] If i.t. physical undecidability of the right kind (which we will discuss shortly) exists, then this determinate ontology of future events may go far beyond the reach of computation. This is not to say that such a picture is necessarily wrong, but only that we might not be able to paint it in full detail—that we might not be able to decide, even given *all* relevant information about the present, whether or not certain events will ever occur.

As Itamar Pitowsky (1996) suggests, some might take such claims as reasons to dismiss such a determinate ontology. Pitowsky points out a parallel between constructivism in mathematics and a certain epistemological conception of determinism. Strong constructivists (such as intuitionists) regard a mathematical sentence as having a

[12] Fine uses similar terminology in [1986] 1996 (pp. 3, 163).

[13] Russell's assertion that the claim is tautological is surprising, since he of all people should recognize the non-trivial import of the definite article. After all, there might not *be* one and only one future, or at least, it might be misleading to adopt a *form of speech* in which the future is determinate. Certainly one cannot infer that there is a unique future from the pseudo-tautological form of the statement in question. However, the point here is not to debunk this "block universe" ontology, but merely to suggest that the world it poses may far exceed our powers of computation.

truth value only if it can in principle be proved or disproved. Pitowsky says that this "in principle" clause is typically cashed out by equating having a truth value with the existence of an effective procedure to determine that truth value (172-173).[14] A similar 'in principle' clause occurs in Laplace's notion of determinism as a kind of predictability by "an intelligence sufficiently vast" ([1814] 1951). Pitowsky takes this too to suggest computability, i.e., that determinism only holds if there is a single Turing machine that can determine any given fact about the future, given sufficiently accurate initial data.[15] Pitowsky then concludes that as a consequence of Moore's undecidability results, a constructivist who denies the truth value of undecidable mathematical propositions must also deny physical determinism:

> The physically significant mathematical structures of theoretical physics are rich enough to allow a translation of many abstract number theoretical propositions into meaningful physical propositions. Therefore, if the physical proposition carries a definite truth value so does its mathematical counterpart. Consequently, if one's intuitions about the reality of space-time and motion are strong, one can take Moore's construction [see above] as a *reductio ad absurdum* of the intuitionist position in mathematics. (1996, 175-176)

---

[14] As formulated by Pitowsky, this supposedly intuitionist view is just a trivial fact. If a particular sentence has a truth value then there is of course an effective procedure to decide the truth of that *particular* sentence: if the sentence is true, for example, a correct decision procedure is just to output 'True.' Decidability is usually understood not as a property of a sentence alone but either (1) of a sentence relative to a *particular* decision procedure or formal theory (e.g., Gödel's famous self-referential sentence is not decided by Peano arithmetic) or (2) of an infinite set (e.g., there is no algorithm to decide which sentences are in the set of theorems of Peano arithmetic). It is not obvious what Pitowsky has in mind, and one of the main points of this section is that in order to pursue arguments like Pitowsky's we must clarify the notions of decidability involved.

[15] Richard Boyd (1972) also takes "predictability in principle" to mean computability by a Turing machine, though like Earman (1986) he denies that determinism should be equated with predictability.

This is not the place for a full evaluation of Pitowsky's argument, but three concerns should be mentioned. Firstly, we have already noted that there are some difficulties in Moore's argument. Secondly, it is not clear that any metaphysical or epistemological claim about mathematics forces any claim about physical reality or vice versa. Brouwer and Heyting, for example, deny that certain mathematical sentences have any truth value precisely *because* of the special nature of mathematical objects, which they regard as mental constructions (Brouwer [1912] 1983, 78; Heyting [1956] 1983, 66, 72). For such intuitionists, the truth or determinacy of mathematical expressions depends on intuitionistic provability alone, not on any correspondence with physical reality.

Finally, Pitowsky's claim, as well as my weaker suggestion that a determinate ontology of the future might outstrip computability, requires the existence of an *actual, particular* system whose i.t. behavior cannot be computed. To establish the existence of such a system would require, among other things, some clarification of the undecidability involved. Suppose for example that several instantiations of a particular universal computing machine actually exist. Then there is no algorithm to determine for *every possible initial state* whether or not one of these machines will "halt," i.e., reach a state designated as an output. But each actual instantiation of such a machine is, at any given moment, in only *one* state. There might be an algorithm to determine whether any of these *actual* systems halt, and if there are only finitely many such systems, then there *certainly* exists such an algorithm. Even if no effective method of deduction can determine the i.t. behavior of a universal machine for *every* possible initial state, it might nonetheless be possible to deduce the i.t. behavior of what few *actual* such systems exist, if any.

Hence, if undecidability is to have much bearing on the predictability of the actual world, it must be a fairly strong kind of undecidability. For example, one might exhibit a system with i.t. behavior that no algorithm can correctly decide over any *neighborhood* of initial conditions. (This will be made more precise in Chapter 3, and in Chapter 4 we will see that it holds for some of Sommerer and Ott's systems (Ott et al. 1994).) To predict the i.t. behavior of such a system, it would not suffice to measure its state up to some neighborhood. One would have to make *full* use of the *exact* initial conditions. One could not deduce the i.t. behavior of such a system from any correct theory, for the initial conditions would have to be expressed by some infinite sequence of symbols, and hence a valid deduction depending on the exact initial conditions would never end. It might be reasonable to say that if a physical system of this kind exists, then a determinate ontology of the infinite future does outreach computability. More to the point, though, this discussion shows that in order to evaluate such claims, we need more precise concepts and results. Furthermore, the suggestion of metaphysical implications provides some philosophical motivation—beyond the usual chore of clarification—to develop such concepts and obtain such results.

1.8. Summary

There is certainly sufficient reason to investigate the i.t. behavior of physical models and the computability or non-computability of such behavior. The history of the question of the stability of the solar system and related questions, the significance they have held for some, and the various partial results obtained, invest the story with considerable human interest. They also illustrate the kinds of interest such problems can have—theoretical, practical, even theological. We have also seen that i.t. predictions are testable in various ways. Further, i.t. undecidability results can guide research away from futile endeavors, such results are not trivially obtained, and it has been suggested that they may themselves have some metaphysical interest.

The pursuit of computability and non-computability results for i.t. problems is part of the broad effort to understand what we can and cannot know and why. For those who are inclined to ground ontology in epistemology, that pursuit may also serve efforts to determine what exists or will exist, in particular whether there is any fact about whether certain events will ever occur. To understand whether or not the computability of i.t. behavior legitimately bears on such questions, or on any questions, rigorous conceptual clarifications will have to be made.

CHAPTER 2

COMPUTABLE SETS OF REAL NUMBERS: A COMPARISON OF CONCEPTS AND MOTIVATIONS

2.1. Introduction

As we have seen (Section 1.6.1), the most obvious notion of a decidable set of real numbers is virtually unsatisfiable; only the empty set and the whole set of real numbers satisfies it. Several authors who were aware of this have introduced more relaxed and useful concepts of decidability for sets of real numbers and sets of points in other continuous spaces. Here we review several such notions, presented informally, and consider their interest, be it practical or theoretical. We also motivate and introduce a new concept, that of decidability up to measure zero. Rigorous definitions are given in Chapter 3.

How should we choose among such notions? According to the approach sketched in the introduction to this dissertation, we should not ask, "What *is* a decidable set of real numbers, *really*?" but rather, "What should we *call* a decidable set of reals?", or to give it even less an illusion of objectivity, "What would we *like* to call a decidable set of reals? To what purposes do we wish to evaluate the decidability of sets, and what concepts (for there may be several) best serve those purposes?"

There are various purposes to consider. One is very general and theoretical: simply to extend the most powerful and aesthetically appealing results of discrete recursion theory to the real numbers. Another is to lay rigorous foundations for the existing theory of numerical analysis, i.e., of practical approximation methods in calculus (Blum, Shub, and Smale 1989; Blum et al. 1998). Some authors are concerned to develop a constructive theory of analysis (e.g., Bishop 1967; Bridges 1999), while others are more concerned with constructivism in physics (e.g. Myrvold 1994, 1995; Pitowsky 1996). For Roger Penrose (1989), the decidability of sets even bears on whether artificial intelligence is possible. There is furthermore a question of logic and epistemology: "What can we learn about the membership of sets of real numbers by systematic reasoning?"

While we will keep such motivations in mind, the ultimate motivating question for this investigation is, "What can and cannot we learn about the behavior of a *physical system* by systematic calculation on the basis of a real-valued model?" For example, when we encounter a planetary system, is it possible to measure its state, and based on this measurement and a mathematical model of the system, to determine whether a planet will ever escape or not? This is a question about the real world, fraught with real-world complications that we will largely ignore except in parts of Chapters 4 and 5. However, as we shop for concepts of decidability, let us remember that we would ultimately like to apply them to this kind of question.

2.2. Decidability, recursiveness, and computability

Before proceeding, it will be useful to redraw a much neglected distinction. Decidability and recursiveness are not the same thing. They are distinct concepts originally introduced for different purposes, and as we will see later, they tend to peel apart when we try to stretch them over the real numbers.

We say a set $A \subseteq \mathbb{N}$ is *recursive* if its characteristic function,

$$\chi_A(n) = \begin{cases} 1 & \text{if } n \in A, \\ 0 & \text{otherwise,} \end{cases}$$

is recursive. The recursive functions on the natural numbers are those built up inductively in certain simple ways from certain very simple functions; the name refers to the familiar notion of a recursive or inductive definition, where the value of, say, $f(n)$ is given in terms of $f(n-1)$ or perhaps in terms of all the values $f(m)$ for $m < n$.[1] The precise definition of recursive function is not important for us at the moment, but it is purely mathematical, involving no reference to computation or effectiveness (Gödel [1934] 1965; for a modern presentation see Soare 1987).

In contrast, the notion of decidability is informal and vague. A set is *decidable* if there is an "effective procedure," an algorithm or recipe, to decide whether or not any

[1] The connection between recursive functions and recursive definitions is discernible, for example, in Skolem 1923, though the functions and predicates discussed there are actually the primitive recursive ones, a subclass of the recursive ones.

given number is an element of the set (Gödel [1934] 1965).

The Church-Turing thesis states that the "effectively calculable" functions on ℕ are precisely the recursive ones (Church [1936] 1965, 100), or equivalently, those computed by Turing machines (Turing [1936-7] 1965). This immediately implies that the decidable *sets* of positive integers are just the recursive ones, and since no clear counterexample to the thesis has been found,[2] "decidable" and "recursive" are now used synonymously.

Yet, in the formative years of computability theory, decidability was an end, and the notion of a recursive set was primarily a means of securing it, a purely mathematical condition guaranteeing decidability. In the lectures where Gödel first publicized the modern notion of recursive set,[3] recursiveness enabled him to prove his incompleteness theorems in virtue of the fact that it *implied* decidability, in particular decidability by a certain axiomatic theory, i.e., provability (Gödel [1934] 1965). However, Gödel did not yet believe the converse, that decidability implied recursiveness as defined in his lectures (Kleene 1986).[4] Even for Church, who proposed recursiveness as a "definition" of

[2] There are abstract constructions of machines that do (in theory) compute non-recursive functions (e.g., Siegelmann 1967), but it is not generally agreed that the operations of such machines constitute effective procedures.

[3] This was based on Herbrand's notion of a recursive function, suggested in a 1931 letter to Gödel shortly before a fatal accident. A related notion of intuitionistic calculability is described in Herbrand [1931] 1967, with suggestive examples of recursively defined functions (618-619).

[4] The lectures were published with a footnote that seemed to anticipate the Church-Turing thesis, but Gödel insisted it did not. In 1965, he wrote to Martin Davis, "The conjecture

computability ([1936] 1965), these two notions were not merely synonymous; if they were, there would have been no need to pose the thesis! Despite Church's "definition," the recursiveness of a set is not the same *concept* as effective decidability but a mathematical idealization of it.

We will avoid confusing decidability and recursiveness here. In order to speak generally about such notions, let us refer to them collectively as notions of a *computable* set. Though 'computable,' like 'decidable,' normally expresses the intuitive notion of effective calculability, we indenture it here as technical jargon for anything vaguely similar to decidability or recursiveness. After all, even the purely mathematical notion of recursiveness expresses a *kind* of computability. Our primary interest, though, is in decidability.

2.3. What is computation?

Since the intuitive notion of decidability involves a notion of effective procedure, or computation, a rigorous conception of decidability on the reals requires a rigorous theory of *computation* on the reals. By this we just mean a theory of effective procedures,

---

stated there only refers to the *equivalence* of 'finite (computation) procedure' and 'recursive procedure'. However, I was, at the time of these lectures, not at all convinced that my concept of recursion comprises all possible recursions" (Kleene 1986; Gödel's emphasis). Apparently, then, he thought computability *coextensive* with some unspecified, all-encompassing notion of recursiveness, but not recursiveness as we know it, and anyway he did not confuse the *concepts* of computability and recursiveness.

whether executed by man or machine.

2.3.1. Recursive analysis

One kind of computation theory on the reals is called *recursive analysis* and is exemplified by the work of Ko and Friedman (1982; Ko 1991) and by Klaus Weihrauch's "Type-2 Theory of Effectivity" or TTE (1985; 2000). In recursive analysis, one regards computation in much the same way as Turing: as the systematic manipulation of finite strings constructed from a finite alphabet.[5] In order to perform computations on real numbers, the numbers must somehow be represented by symbol strings. Since there are not enough finite strings to represent all the real numbers, reals are represented by infinite strings, or equivalently, by infinite sequences of finite strings, or by infinite sequences of natural numbers. For the sake of simplicity, we choose the latter here; real numbers will be coded as infinite strings of *natural numbers*. Our theory will nonetheless be entirely equivalent to one using infinite sequences of finite strings from a finite alphabet, since all such finite strings can be effectively enumerated and are therefore interchangeable with the natural numbers.

[5] A related approach, which sometimes called *constructive* analysis, is similar but concerned only with the field of *computable* real numbers in the sense of Turing [1936-7] 1965, e.g., with computable sets of computable reals and computable functions on the computable reals (Mazur 1963; Bishop 1967; Aberth 1980). The approach we are calling recursive analysis, and which we adopt throughout most of this essay, treats the entire continuum of real numbers (and other continua), and the computability of function values *relative* to given arguments, which might themselves be non-computable.

A *representation* of the set $\mathbb{R}$ of real numbers, then, is a map from a set $A \subseteq \mathbb{N}^{\mathbb{N}}$ onto $\mathbb{R}$. (See Definition 3.2.2.) For example, the familiar decimal representation maps a sequence of natural numbers to a given real number if the sequence is just a decimal expansion of the real number. (For details, see Definition 3.2.3(ii).) However, this representation turns out to be inconvenient for a careful and broad study of computability, for reasons that will be explained in the next sub-section.

A more useful representation, which we will call the *standard* one, codes a real number x as, in effect, a list of all open intervals that have rational endpoints and contain x. In Section 3.2.1 we fix an effective enumeration I of the open intervals with rational endpoints. The standard representation of the reals maps a sequence ϕ of naturals to a real number x if and only if the numbers in ϕ correspond, via I, to all and only those rational intervals containing x (Definition 3.2.3(i)). To generalize this to $\mathbb{R}^n$, we call a product

$$(q_1,q_2)\times(q_3,q_4)\times\ldots\times(q_{2n-1},q_{2n})$$

of intervals, where each q_i is a rational number, an *open rational n-interval*. We fix an effective enumeration I^n of these, too, in Section 3.2.1. The *standard representation* of $\mathbb{R}^n$ then maps a sequence ϕ of natural numbers to a point x if and only if the numbers in ϕ correspond via I^n to precisely those open rational n-intervals containing x (Definition 3.2.3 (i)). In that case, we will call ϕ a *standard name* for x.

In recursive analysis, the systematic nature of computation is usually modeled by an abstract computing machine, namely the Turing machine, with special conventions for infinite inputs and outputs[6] and perhaps some inessential modifications. Ko (1991) and Weihrauch (2000) specify such machines in full detail. Here we adopt a model similar to Ko's: a two-tape Turing machine, where one tape holds the infinite input, and perhaps a separate finite input as well,[7] while scratch work and output are written on the other tape.[8] We do not specify the full details of such machines but instead rely on a version of the Church-Turing Thesis extended to functions on infinite arguments:

> **Generalized Church-Turing Thesis for infinite arguments:** Of the functions that map infinite sequences of natural numbers to individual naturals, those that can be computed by an effective procedure in finitely many steps are precisely those that can be computed by a two-tape Turing machine where one tape supplies the infinite input.

This is only slightly stronger than the Generalized Church-Turing Thesis used in the study of r.e. degrees (Soare 1987), which states that the functions that map *individual* natural

[6] Even Turing imagined his machines to generate infinite output ([1936-7] 1965), but in many modern discussions this is not permitted.

[7] Of course, a finite input and an infinite one can be coded together as a single infinite input, but it will often be useful to distinguish a separate finite input.

[8] This is similar to an *oracle Turing machine,* and the infinite input might be called an *oracle* as in Ko (1991). This terminology can be confusing, though, for in our use of such machines, the "oracle" is not an *aid* to computation, answering non-computable questions as in the theory of r.e. degrees (Soare 1987); it merely supplies the *input.* Though our machines of course make use of their "oracles," there is no worry that this will make our model of computation too powerful, for we are only concerned with what *function* a machine computes over a broad *domain* of oracles. Any one oracle input to a machine has an effect only on the one computation taking that oracle as input, not on the *function* computed by the machine, taking *each* oracle to some output.

numbers to natural numbers and can be computed by some algorithm, using reference data supplied by a *fixed* infinite sequence ϕ, are precisely those computed by some two-tape Turing machine supplied with that *same* fixed sequence ϕ. (See note 8.) ere we merely allow ϕ to vary and thus obtain functions on infinite arguments. The use of a two-tape machine is inessential, for the same computations could be carried out by an ordinary one-tape Turing machine, given appropriate input and output conventions. As the thesis indicates, such details make no difference for our purposes.[9]

Like Ko, we permit infinite inputs but not infinite *outputs*. Instead we define computability of a real-valued function in terms of finite computations of approximations. We say that a function $f\colon \mathbb{R}^n \to \mathbb{R}^m$ is *computable* if there is a machine M that, given a standard name of x and a natural number i, will output the i^{th} symbol of a particular standard name for $f(x)$—that is, if $\{M(x, i)\}_{i \in \mathbb{N}}$ is a standard name for $f(x)$ (Definition 3.2.5).[10] Below we discuss notions of computable set in the same framework.

[9] One exception: Weihrauch permits his "Type-2 machines" to output infinite strings, with the restriction that the head on an output tape can only move from left to right. The latter restriction is significant; without it, Weihrauch's machines would be more powerful than Ko's, and in particular, the step function discussed below could be computed by such a machine. However, in attempting to use such a machine, one might never know whether a given output symbol were really part of the final output, or whether the machine would at some point come back and revise it. In that case, the machine would not provide any reliable information, even of an approximate kind, in finite time, so this would not be a very useful sort of computation.

[10] This is a generalization of Grzegorczyk's notion of a computable function on the real numbers (1955, 1957).

2.3.2. The choice of representation

We are now better equipped to explain why we do not usually use the familiar decimal or base-b digital representation of reals in recursive analysis. If we replace the standard names in the definition of computable function in the preceding paragraph with base-b digital names, it turns out that even the basic operations of addition and multiplication are non-computable (Proposition 3.2.10).[11] This may be puzzling, since we in fact add and multiply using digital expansions on a daily basis. What is meant here is that there is no general algorithm to find the n^{th} digit of $x + y$ or xy in finite time, due to the fact that numbers arbitrarily close to each other sometimes have very different base-b expansions.[12] (The proof of 3.2.10 is instructive.) In our usual applications we are not concerned with computing particular digits correctly; we are content if we can produce decimals that are arbitrarily close to, say, $x + y$ in the usual Euclidean metric on $\mathbb{R}$, even if the digits of these approximations are not the same as those of the exact sum. In effect, what we really want—and have—is access to a *sequence of approximations* that converge to $x + y$ in the Euclidean metric and in an effective way. This is just what we

---

[11] To state this precisely, we must extend base-b digital representations to $\mathbb{R}^2$. That is easy enough, but it is even easier to show as an example that the function $f(x) = x + 2/3$ is not computable with respect to the decimal representation (Proposition 3.2.10).

[12] In Weihrauch's language (2000, 68), digital representations are not *admissible,* meaning that digital names and standard names cannot be translated into one another continuously (with respect to the Baire topology, or if one adopts a finite alphabet, as Weihrauch does, with respect to the Cantor topology). This is just because there is no continuous function taking each $x \in \mathbb{R}$ to a base-b name for x.

call a regular Cauchy representation (Definition 3.2.3(iii, iv)), and it is equivalent to the standard representation (Propostion 3.2.9). Since the standard and Cauchy representations seem to capture this intuition about what sort of approximative computability we would like, and since they result in a more reasonable notion of computable function, they are preferred in recursive analysis.

With regard to notions of *decidability* or *near-decidability*, construed in a certain strict way, the choice between these representations is not critical. If there is no algorithm to determine whether or not an arbitrary real number x is in the set A, then we would like an algorithm that does so for, in some sense, *most* real inputs. That is to say, we would like to be able to compute the characteristic function of A, which is a function into $\{0, 1\}$, everywhere except perhaps on some "small" set of inputs. For the computation of functions from subsets of $\mathbb{R}$ into $\{0, 1\}$, the digital, standard, and Cauchy representations are all equivalent (Proposition 3.2.14). However, some of the notions of recursive set that we will consider below involve other properties, such as the computability of the distance between a point x and the set A. In order to evaluate such properties appropriately, we adopt the standard representation. Also, the standard representation generalizes more easily to other topological spaces (Chapter 3).

2.3.3. Recursive analysis versus real RAMs

There is an ongoing debate between proponents of recursive analysis, based on discrete symbolic representations of real numbers, and those who treat real numbers themselves, regarded as abstract entities independent of any representation or construction, as direct objects of computation. This is done by means of a generalization of the so-called random access machine (Shepherdson and Sturgis 1963) called a real-valued random access machine, or real RAM (Preparata and Shamos 1985). Here we are mainly concerned with the recursive analysis approach, but let us now contrast its virtues with those of the real RAM model.[13]

One influential real RAM study is that of Blum, Shub, and Smale (1989; Blum et al. 1998). Their machine is described as a flow chart where each node is associated with a computation step. (See Figure 2.1.) In a single step, this machine can do one of two things: (i) It can update the values of any number of variables according to prescribed rational functions (functions each defined by a ratio of polynomials[14]); for example, it can reset the value of a pair of variables (x, y) to $(g_1(x, y), g_2(x, y))$, where g_1 and g_2 are rational functions. Or (ii) it can perform an exact comparison and branch accordingly; that is, it can answer a question such as '$h(x, y) < 0$?' where h is a rational function, so that the answer determines to which node the computation will proceed. (Note that one can combine two of these comparison nodes to obtain an equality test.) A computation

[13] There are several other approaches to computable analysis, many of which are closely related to recursive analysis. See Weihrauch 2000 (249-268) for a survey.

[14] Blum et al. define these machines to take arbitrary rings as their domains, so if the ring in question is not a field, only polynomial functions, rather than rational functions, are permitted. Here we are mainly concerned with the fields $\mathbb{R}^n$.

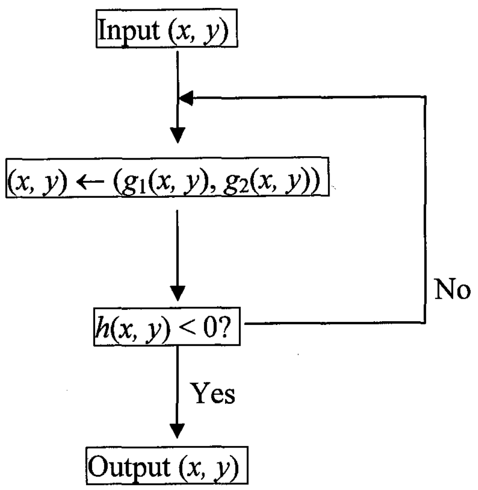

Figure 2.1. A simple real RAM (Blum et al. 1998). Here g_1, g_2, and h are rational functions.

is completed when it reaches a designated output node on the flowchart.[15]

The most pertinent virtue of the real RAM model for us is that it directly suggests a non-trivial notion of decidable set. Say a subset A of $\mathbb{R}^n$ is *real RAM decidable* if its characteristic function is real RAM computable, i.e., if there is some real RAM that, given input $x \in \mathbb{R}^n$, halts in finite time with output

$$\chi_A(x) = \begin{cases} 1 & \text{if } x \in A, \\ 0 & \text{otherwise.} \end{cases}$$

[15] For precise definitions of and rigorous results on these machines, see Blum, Shub, and Smale 1989 and Blum et al. 1998.

This notion is not trivial or empty. The interval [0, 1], for example, is real RAM decidable, while the middle-third Cantor set is not (Blum et al. 1998; see Proposition 3.6.3 below for construction of a Cantor set.)

On the face of it, the fact that simple sets like [0, 1], [0, ∞), and the unit disk in $\mathbb{R}^2$ are real RAM decidable speaks well of that concept; if any non-trivial sets should be considered computable, surely these are among them.[16] Also, real RAM decidability bears some analogies with the notion of a decidable set in $\mathbb{N}$. It is similarly symmetric, both in the sense that a set is real RAM decidable if and only if its complement is, and in that it implies that some real RAM will tell us if a point x is in a set A and will also tell us if x is *not* in A. Furthermore, the recursive sets of *integers* are precisely those sets of integers that can be decided by a real RAM defined with integer variables and coefficients (Blum, Shub, and Smale 1989, 33). In this sense, real RAM decidability is a formal generalization of classical decidability on $\mathbb{N}$.

However, in a number of respects, real RAM decidability is unsatisfying as a model of computation as usually conceived. In one way it seems a little too strong, for a relatively simple set like the closed epigraph $E = \{(x, y): y \geq e^x\}$ of the exponential function is not real RAM decidable; in fact, no graph, closed epigraph, or closed

---

[16] Penrose (1989) suggests that the unit disk and also the closed epigraph $E = \{(x, y): y \geq e^x\}$ of the exponential function ought to come out recursive. Earman (1986, 119) proposes that a simple step function such as that defined below ought to be considered computable, which suggests that sets like [0, ∞) should be considered decidable. See Section 2.4.

hypograph of an everywhere transcendental and continuous function f is real RAM decidable (Brattka 1993b). Yet E and many other such sets seem intuitively tame and computable, and Roger Penrose (1989) suggests that sets like E ought to be considered recursive. (Penrose also asks whether the Mandelbrot set is recursive. Blum and Smale (1993) answer that it is not real RAM decidable, but since real RAM decidability does not satisfy all of Penrose's criteria, this answer is not completely appropriate. See Brattka 1993b for a full discussion of this point.)

There are several other ways in which the real RAM seems an ill-fitting model of computation. Most significantly, a real RAM can perform exact comparisons between real numbers. A Turing machine (even with infinite inputs) cannot do this; given two names representing real numbers, it can eventually tell us if they are *not* equal, but it can never compare all the symbols of the two names, so it can never be "sure" that they *are* equal. This is closely related to the fact that Grzegorczyk computable functions are continuous (Grzegorczyk 1955). Hence, not even the step function

$$f(x) = \begin{cases} 0 & \text{if } x < 0, \\ 1 & \text{otherwise,} \end{cases}$$

is Grzegorczyk computable, but it is real RAM computable. This is a significant difference, for as we will see in Section 2.4, if a Turing machine could compute such a function, it could also solve the famously unsolvable halting problem. In fact, allowing a machine to evaluate comparisons exactly in finite time has been shown to be, in a sense, exactly equivalent to a "jump," i.e., to a solution of the halting problem (Boldi and Vigna 1999).

The fact that real RAMs are in this respect stronger than Turing machines is in itself no blemish, but the real RAMs' exact comparisons do not seem to correspond to anything in actual computational practice. In pure mathematics, the real numbers on which we actually perform computations are specified by definitions, not given as objects in themselves that immediately reveal their differences with other numbers. We can sometimes generate effectively converging approximations to a given number, but this does not provide us any way to decide when two such approximations are equal. Borel made this point as early as 1912, working from an informal, intuitive notion of computability. He writes about exact comparison,

> One is not absolutely sure, from a theoretical point of view, that it does not present insoluble difficulties, because one can conceive of two numbers such as the following:
>
> $$\left(\int_{-\infty}^{\infty} e^{-x^2} \mathrm{d}x\right)^2, \quad \int_{-\infty}^{\infty} \frac{\mathrm{d}x}{1+\mathrm{x}^2}$$
>
> which are identical up to any number of decimal places, and nevertheless one does not know how to prove equality. (163-164, quoted in Myrvold 1994, 58)

Borel's recognition of this difficulty supports the view that it is not an artifact of any particular technical model of computation but inherent in the intuitive notion.[17]

[17] Thanks to Wayne Myrvold for making this point about Borel, both in his dissertation (1994) and in his comments on a draft of this chapter.

In applications to nature, real numbers are given by measurements, which also are only approximations. At best we might be able to carry out a sequence of increasingly accurate measurements of a single quantity, but again this is not enough to enable exact comparisons.

One might think of the real RAM as a model of analog computation, where a real number is not represented by discrete symbols but by some physical quantity, such as the amount of current in a circuit, that is analogous (e.g., proportional) to the real number represented. Even in this context, though, exact comparisons are suspect. Such comparisons would again constitute perfectly accurate measurements of physical quantities.[18] Furthermore, it seems that the construction of an analog machine capable of exact comparisons would require perfect measurements. In contrast, Turing machines (with limited tape space) can actually be constructed with finite measurement skills. In fact, any Turing machine can be simulated by a common electronic computer, except only that the computations of actual digital machines are limited by time, energy, and storage space. On the other hand, even a limited real RAM cannot be implemented in any obvious way until we figure out how to make perfect measurements.

[18] Aside from apparent practical limitations on measurement, which might in general be surmountable, quantum mechanics suggests that there are *fundamental* obstacles to perfect measurement inherent in the very structure of the world. However, this objection applies as well to an assumption that we *will* make throughout most of this investigation, namely that one can make measurements of *arbitrarily fine* but *imperfect* accuracy. Except for a few remarks later on, we will largely ignore quantum considerations.

In any case, we will exclude analog computation from the scope of these investigations. We will not say that analog computation is not computation; analog computing machines were discussed as early as 1876 and built as early as 1931 (Campagnolo, Moore, and Costa 1998), pre-dating the electronic digital computer, the hypothetical Turing machine, and classical discrete recursion theory. However, analog computation should not be confused with the human activity that dominates mathematical practice. Turing's is a theory of *human* computation. There it is not enough that some sort of machine could compute a given function in some sense; it has to be executable by a *person*, or a finite team of people, with finite measurement skills.[19] Turing's "machine" is intended as a model of such human activity. In practice, whether using paper and pencil or electronic machines, most people still compute with discrete symbols. Hence the powers and limitations of *these* activities are especially worth investigating.

Of course, Blum and company regard the real RAM as a model of actual practice too, in particular of the practice they call "scientific" computation. But such computation is done with digital computers, using only rational approximations to real numbers! Hence it must be admitted that the real RAM is at best a loose model that ignores the effects of digital implementation. (See Weihrauch 2000, 262-264 for some details of these effects.) Certainly we can refine this model, and we can even eliminate the exact

[19] This is why Turing does not permit infinitely many distinct symbols; "there would be symbols differing to an arbitrarily small extent," implying that we could not then discern them ([1936-7] 1965, 135).

comparisons as well as the other troubling features that we have not discussed, but the more we do so, the closer we come back to the Turing machine. In fact, Brattka and Hertling (1996) have introduced a modified "feasible" real RAM, which still treats abstract real numbers as the direct objects of computation, without the mediation of symbolic representations, but which does not employ exact comparisons. The real functions computable by these feasible real RAMs are precisely those computable by a Turing machine with infinite input.

One might object that permitting infinite inputs, as we do in recursive analysis, is much like assuming perfect measurement accuracy. However, we require a Turing machine to provide each approximate output in finite time, so each such output depends on only a finite portion of the input; this is called the Use Principle (Soare 1986). (See Section 2.4 and Theorem 3.2.16 below.) Hence what we are really assuming is access to *arbitrarily many* input symbols, and thus when these symbols approximate some real number, we are assuming access to *arbitrarily* accurate but never quite *perfect* approximations to the input. This is perhaps only a slightly more plausible fantasy than perfect measurement, but it can hardly be avoided. Were we to assume any particular finite limit on measurement accuracy, we would not have a sufficiently general theory. We would live in perpetual fear that actual measurements might surpass the limits assumed by our model. Anyway, even if there are definite limits on physical measurement, these limits depend on the unit of measure chosen. Similarly, if we assumed a particular length limit on inputs, our model could eventually be surpassed by actual machines or even teams of human beings. By allowing infinite inputs and assuming that arbitrarily accurate measurements are possible, we abstract away from the specific limitations of particular computers and measurement techniques. Thus we can show, for example, what functions cannot be computed *no matter how accurate* the data are.

Throughout the remainder of this essay, therefore, 'machine' will refer to some equivalent of a Turing machine permitting infinite input and giving finite output, as discussed in Section 2.3.1.

2.4. In defense of naïve decidability

Even within the symbolic, approximative framework of recursive analysis, it is not that there is *no* especially natural and obvious notion of a decidable set of real numbers. In fact there is one: we could say a set $A \subseteq \mathbb{R}^n$ is decidable if there is some algorithm or Turing machine that, given *any* $x \in \mathbb{R}^n$, will determine in finite time whether $x \in A$ or not. Let us call this *naïve decidability*. The problem is that on the discrete conception of computation, this notion is virtually unsatisfiable. Even a simple set like the interval $[0, \infty)$ is naïvely undecidable, as is the closed epigraph of e^x, as well as any set of naturals embedded in $\mathbb{R}$, whether recursive in the classical sense or not.

A helpful way of seeing this is in terms of the Use Principle from the theory of r.e. degrees (Soare 1987). This states that if a Turing machine is supplied an infinite string of information and it eventually halts, it does so having *used* only a finite segment of that information (Theorem 3.2.16). The infinite inputs that concern us are, in effect, approximations to a point, so it follows from the Use Principle that a machine can only make use of *inexact* information about that point. Thus we have what we might call the...

> **Topological Use Principle**[20] **(TUP**; Proposition 3.2.17): If a machine M halts on input $x \in \mathbb{R}^n$ with output q, then it does so for all inputs y in a neighborhood of x.

This principle will lead us to many undecidability results. It implies that at the boundary of a set A, where every neighborhood contains elements of both A and its complement, no machine can correctly decide membership in A over a neighborhood. If a machine classifies a boundary point as belonging to A, for example, then it will also do so for many nearby points that are not in A. (See Figure 2.2.) Hence every set with a non-empty boundary is naïvely undecidable.

Some seem to regard the naïve undecidability of such simple sets as $[0, \infty)$, or equivalently, the non-computability of the step function defined in Section 2.3.2, as an unfortunate technical artifact of recursive analysis that could perhaps be removed (e.g., Earman 1986). However, such results are more fundamental than they might appear. As

[20] Note that under the standard representation, a number has many names. In this statement of the TUP we have ignored details about how M responds to different standard names for x. Those detail are spelled out in Proposition 3.2.17.

Myrvold shows, an algorithm to compute the step function (in the symbolic framework of recursive analysis) could be used to solve the famously unsolvable halting problem of classical recursion theory (1997). Likewise, a foolproof decision procedure for $[0, \infty)$ would constitute a solution to the halting problem. Both facts can be proved as follows: Let $\{\varphi_i\}$ be an effective enumeration of the partial recursive functions on $\mathbb{N}$. For each i and j, we effectively construct a name for a number $n(i, j)$ that is equal to zero if $\varphi_i(j)$ is undefined but less than zero if φ_i is defined on j. (See Myrvold 1997 for details of the construction.) Then, if we have a procedure to compute arbitrarily accurate

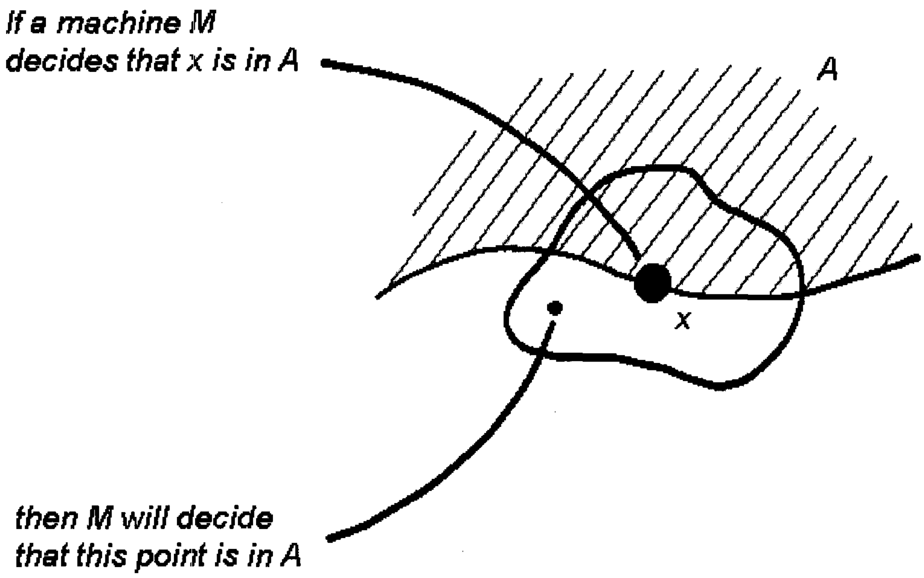

Figure 2.2. Boundary points and the TUP. If a machine (or procedure) halts on an input point x, then it will give the same output for all points in some neighborhood of x. Therefore, if x is a boundary point of a set A, such a machine will incorrectly decide A at many points near x.

Perhaps it is not so strange that only trivial sets of reals are strictly decidable. This fact bears at least a superficial resemblance to an important result in classical recursion theory called Rice's theorem. An *index set* is a set of the form $\{i \in \mathbb{N}: P(\varphi_i)\}$ (with $\{\varphi_i\}$ as above), defined in terms of some *extensional* property P of a function φ_i, such as the property of being total, having a non-empty domain, being one-to-one, *et cetera*—any property of the *function* that does not depend on the particular way it is defined or represented. Rice's theorem says that the only recursive index sets are the trivial ones: the null set and $\mathbb{N}$ (Soare 1987, 21). So only trivial properties of partial recursive functions are decidable from their integer names, just as only trivial properties of real points are decidable from their string names. The connection may not be superficial, since these string names are sequences, i.e., functions on $\mathbb{N}$. We will not attempt to develop the connection further, though; we only note that the situation with naïve decidability is not so unfamiliar and so perhaps not so troubling.

Naïve decidability is in any case a very natural notion of decidability for sets of real numbers. In its definition it comes closer to the classical notion of decidability over $\mathbb{N}$ than any of the relaxed notions we will consider below. It also bears the right kinds of symmetry with respect to complementation, and it is equivalent to the computability of a characteristic function. Given all of the above observations, it would not seem unreasonable to take naïve decidability as *the* natural generalization of decidability to sets of real numbers, and bravely face the fact that non-trivial sets of real numbers are just not decidable.

Nonetheless, some sets are even less decidable than others. Those who have claimed that certain models of physical systems have undecidable behavior have given sophisticated arguments that suggest something more than this trivial, naïve undecidability. Even from a purely mathematical point of view, some sets are non-computable in ways that go beyond mere non-triviality. We therefore proceed to consider notions of computable set that attempt to capture such intuitions.

2.5. Recursive open and closed sets

Perhaps the two most popular notions of a computable set of reals, at least within the recursive analysis approach, are that of a recursive closed set and the complementary notion of a recursive open set. To explain these notions, we first introduce r.e. ("recursively enumerable") open and closed sets.

The concept of an r.e. open set is especially fundamental to our investigation, and all of the notions of effectiveness to be discussed below are closely related to it. A set $A \subseteq \mathbb{R}^n$ is *r.e. open* if there is some machine that halts on an input $x \in \mathbb{R}^n$ if and only if $x \in A$ (Definition 3.3.1).[21] This definition straightforwardly generalizes the notion of an r.e. subset of $\mathbb{N}$. By the TUP, these are indeed open sets (Remark 3.3.4). What makes them so fundamental is the fact (obvious from the definition) that the domain of any

[21] These were called recursively open sets in Lacombe 1957, 1958.

computable function from a set of real numbers into $\{0, 1\}$ (or $\mathbb{N}$) is an r.e. open set. In this respect too, the r.e. open sets are analogous to the r.e. subsets of $\mathbb{N}$. (The domains of computable functions into $\mathbb{R}^n$ are more complex; see Weihrauch 2000, 31).

On the other hand, an *r.e. closed* set in $\mathbb{R}^n$ is a closed set such that we can effectively list all the open rational n-intervals that intersect it (Definition 3.3.7; cf. Weihrauch 2000). Equivalently, it can be defined as a closed set with a dense computable sequence (Zhou 1996).

Since the complement of an open set is closed and vice-versa, it seems fairly natural to define a *recursive open* set as an r.e. open set with an r.e. closed complement, and a *recursive closed set* as an r.e. closed set with an r.e. open complement (Definition 3.3.8), as do Weihrauch and others (Kreitz and Weihrauch [1984] 1987; Zhou 1996; Brattka and Weihrauch 1999; Weihrauch 2000; Brattka 2003a, b). Let us also call a bounded recursive closed set a *recursive compact set.* Since these concepts are often used within Weihrauch's Type-2 Theory of Effectivity (TTE), we say a recursive open or recursive closed set is *TTE recursive.* (See Definition 3.3.8 (iii)).

Superficially, at least, TTE recursiveness has just about all the properties one would like in a notion of a computable set of reals. An open or closed set is TTE recursive if and only if its complement is, which holds if and only if both sets are r.e. Simple sets like the interval $[0, \infty)$, the closed unit disk in $\mathbb{R}^2$, and the closed epigraph of e^x turn out to be recursive. Any set $A \subseteq \mathbb{N}$ is recursive in the classical sense if and only if it is a recursive closed set $A \subseteq \mathbb{R}$ (Weihrauch 2000, 126). Also, just as a function on $\mathbb{N}$ is recursive if and only if its graph is recursive, a continuous function on $\mathbb{R}$ or on a recursive compact subset of $\mathbb{R}$ is recursive if and only if its graph is a recursive closed set (Zhou 1996; Weihrauch 2000; see Brattka 1993a for a thorough discussion). Furthermore, as we will discuss below, TTE recursiveness is equivalent to the computability of a generalization of the characteristic function.

Another nice feature of TTE recursiveness is the fact that recursive open and closed sets can be effectively graphed to arbitrary precision, in a certain sense (Weihrauch 2000; cf. Brattka 2003a). Suppose for example that A is a recursive closed subset of the unit square in $\mathbb{R}^2$. Let us divide the unit square into k^2 disjoint squares of equal size, each containing a colored pixel. By the definitions of recursive open and closed sets, there is an effective procedure to color each pixel either black or white in such a way that every white pixel will lie entirely outside A, and every black pixel will lie entirely within a small distance, inversely proportional to k, of the set A. (See Figure 2.3.) Hence by choosing sufficiently large k, a plot of A can be made arbitrarily accurate with respect to *distance*. This is clearly a useful property. We will consider it more critically in Section 2.7.

Unfortunately, TTE recursiveness has some troubling asymmetries. By definition, an open or closed set is recursive if and only if it is r.e. and its complement is r.e. However, the sense of 'r.e.' is different for open and closed sets.

The notion of r.e. open is analogous to classical recursive enumerability, especially with regard to decidability. The r.e. open sets are precisely the semi-decidablesets, in the sense that there is an algorithm that will tell us that a given x is in A if and only if x really is in A. Furthermore, the r.e. open sets are recursively enumerable in a literal sense; they are precisely those sets that are unions of recursive sequences of open rational n-intervals (Proposition 3.3.3). One might say we can enumerate them in clumps.

Though the notion of r.e. closed also has nice properties, it is not nearly so analogous to classical recursive enumerability. There is *never* an algorithm that halts on all and only the elements of an r.e. closed set; such semi-decidable sets are always open. Nor is an r.e. closed set ever the set of inputs on which some machine halts, and nor are r.e. closed sets recursively enumerable in an intuitive sense. As noted, we can enumerate the points in a dense subset of any r.e. closed set, and by definition we can enumerate the

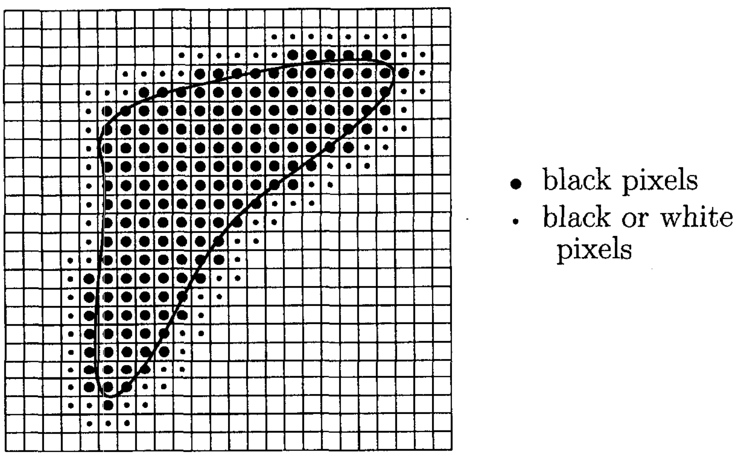

Figure 2.3. Pixelated graph of a recursive closed set (Weihrauch 2000).

open rational intervals that *intersect* it, but we cannot enumerate an entire r.e. closed set, not even in clumps as with the r.e. open sets. In fact, Hemmerling (2003) exhibits an r.e. (and further, recursive) closed set A such that it is impossible to decide which *naturals* lie in A, even though no naturals lie on the boundary of A. In short, r.e. closed means we can determine which points are *near* the set, but not which points are *in* it.

Thus the apparent symmetry of TTE recursiveness suggested by its relation to the notions of r.e. open and r.e. closed is somewhat illusory. Likewise, the other apparent symmetry—that an open set is recursive if and only if its closed complement is recursive—also runs rather shallow, since these notions of recursive are so different. Recursive open sets are at least intuitively r.e., while recursive closed sets are much less so.

As mentioned, TTE recursiveness is equivalent to the computability of a certain generalization of a characteristic function. That generalization is the distance function $d_A(x) = \inf_{y \in A} \|y - x\|$ (Weihrauch 2000, Brattka 2003b); a closed set is recursive if and only if its distance function is computable (Weihrauch 2000). As Brattka points out (op. cit.), the characteristic function of a subset of $\mathbb{N}$ is equal to the distance function under the discrete metric on $\mathbb{N}$. In this respect, the distance function is a formal generalization of the notion of a characteristic function.

Yet, for reasons connected with those considered just above, the appropriateness of this generalization is questionable. Firstly, the equivalence of recursiveness with computability of the distance function applies only to closed sets. There are many sets *neither open nor closed* that have computable distance functions. Hence it is not true that an *arbitrary* set A is TTE-recursive if and only if the distance function is computable. Secondly, the distance function is distinctly asymmetric with regard to sets and their complements. It takes various positive values in the exterior of a set, but it is constantly zero within the set. (See Section 3.1 for topological concepts such as *interior*, *exterior*, and *boundary*.) Consequently, the computability of d_A immediately implies that the exterior of A is semi-decidable: to determine that $x \in \text{ext}(A)$ (if x *really is* in $\text{ext}(A)$), just compute the distance function to sufficient accuracy to be sure that it is non-zero. Yet d_A offers no such information about A itself or the interior of A. No matter how accurately we compute $d_A(x)$ we can never be sure that it is *exactly* zero, so we can never be sure that a point x is in A on that basis.

This is hardly surprising, since we have already noted that closed sets are not semi-decidable, but then how useful is this generalization of the characteristic function? The original motivation for defining the notion of a recursive set in terms of the characteristic function was to establish a mathematical condition for *decidability*. Immediately after defining 'recursive relation (or class),' Gödel explains,

> [R]ecursive relations (classes) are decidable in the sense that, for each given n-tuple of natural numbers, it can be determined by a finite procedure whether the relation holds or does not hold (the number belongs to the class or not), since the representing function is computable. ([1934] 1965, 44)[22]

[22] The notion of *recursive function* referred to in this quote is what we now call *primitive recursive*. Gödel introduces the modern notion of recursive function ("general recursive") later

The same motivations are clear in the paper where Church first states his famous thesis. Church gives motivating problems in the introduction, such as, "to find a *means of determining* of any given positive integer n whether *or not* there exist positive integers x, y, z, such that $x^n + y^n = z^n$" ([1936] 1965, my emphasis). A recursive distance function, on the other hand, does not serve as a guarantee of decidability, and even the promoters of TTE recursiveness recognize this. Brattka writes, "Although recursiveness of subsets of Euclidean space in this sense does not correspond to the intuition of 'decidability', it is a formal generalization of the classical notion of recursiveness" (Brattka 2003b). We will soon see that there is another continuous formal generalization of the characteristic function that implies something much closer to decidability.

2.6. Strong recursiveness, Δ-decidability, and decidability ignoring boundaries

2.6.1. Strong recursiveness

A natural strengthening of TTE recursiveness rectifies the asymmetries pointed out in the preceding section. Following Zhou (1996), we say a closed set A is *strongly recursive* if A is recursive closed and its interior, int(A), is recursive open. Similarly, an open set A is *strongly recursive* if A and its closure cl(A) are both recursive. Hence, any

---

in the lectures. Regardless, the significance of a (primitive) recursive characteristic function as a guarantee of decidability is clear here.

strongly recursive set has a recursive open interior and exterior. (Hertling calls these sets *birecursive* in 1999.)

Like TTE recursiveness, this concept seems to have just about everything, but more so. The intuitively simple sets we have discussed—$[0, \infty)$, the unit disk, the epigraph of e^x, recursive sets of naturals embedded in $\mathbb{R}$—are all strongly recursive. Also, the correspondence between the recursiveness of a function and that of its graph is not lost; the graph of a continuous function f on $\mathbb{R}$ or on a recursive compact subset is not only recursive but strongly recursive if and only if f is computable.[23] Clearly strong recursiveness also implies the existence of an arbitrarily (distance-wise) accurate graphing procedure, since it implies TTE recursiveness, though it is stronger than necessary for this purpose.

The symmetry of strong recursiveness with respect to complementation runs deeper than that of TTE recursiveness. We could characterize the former in terms of a strengthened semi-recursiveness: say an open or closed set A is *strongly r.e.* if its closure and its interior are both r.e.[24] It follows immediately that an open or closed set A is strongly recursive if and only if A and A^{C} are strongly r.e. Also, an open or closed set is

[23] This is because the graph of a continuous function on a closed set is contained in its own boundary, and as Hemmerling notes (2003), such a set is recursive if and only if it is strongly recursive.

[24] In fact, the closure of an r.e. open set in $\mathbb{R}^n$ is always r.e. closed, but the interior of an r.e. closed set is not necessarily r.e. open (Hemmerling 2003).

strongly recursive if and only if its complement is, but more significantly for our purposes, this implies that both the interior and exterior are r.e. open. Hence, though a strongly recursive closed set is still not even semi-decidable or recursively enumerable in the same full and intuitive senses as an r.e. open set, at least its interior and exterior are.

Strong recursiveness is also equivalent to a generalization of the characteristic function, but one that better serves the cause of decidability than the distance function. Let the *symmetric distance function* $\Delta_A(x) = d_A(x) - d_{A^C}(x)$, where A^C is the complement of A. An open or closed set A is strongly recursive if and only if Δ_A is computable (Hemmerling 2003). This function is perhaps not a very *direct* formal generalization of the characteristic function, but if again we impose the discrete metric on $\mathbb{N}$, then for any subset A of $\mathbb{N}$, $\chi_A(n) = (1 - \Delta_A(n))/2$. (For example, if $n \in A$, then $d_A(n) = 0$ and $d_{A^C}(n) = 1$, so $\Delta_A(n) = -1$ and $(1 - \Delta_A(n))/2 = 1 = \chi_A(n)$.) In any case, the value of the symmetric distance function clearly indicates whether or not a point x is in A except in borderline cases. For any x not on the boundary of A, $\Delta_A(x)$ is non-zero, and in that case we can determine whether it is positive or negative by computing it to sufficient accuracy. In this respect the computability of Δ_A guarantees something much more like decidability than does that of the one-sided distance function. All in all, strong recursiveness is more like a notion of decidability than TTE recursiveness is.

2.6.2. Δ-decidability and decidability ignoring boundaries

So far we have considered only notions of computable open and closed sets. Studies adopting the TTE approach typically consider only special classes of sets, such as the open and closed sets, regular sets, or closed convex sets (Ziegler 2002, Kummer and Shäfer 1995). This approach has definite advantages. For one, it limits the cardinality of the class of sets under consideration to 2^{ω}, so that every set can be represented by a countably infinite string. The computability of open and closed sets then reduces to the computability of corresponding strings.

However, Hemmerling has pointed out that if the intervals (0, 1) and [0, 1] are recursive, surely the half-open interval [0, 1), being hardly more complicated than the other two, should also be considered recursive. Similarly, the embedding $\{0\} \times (0, 1)$ of the open interval (0, 1) into $\mathbb{R}^2$ ought to come out recursive. These sets are neither open nor closed, and therefore they are not TTE recursive, let alone strongly recursive. Furthermore, in applications one may encounter sets that are not open, closed, convex, or regular, or one may simply not know whether a given set has such properties. Hence it seems useful to discuss the computability of arbitrary sets in $\mathbb{R}^n$, as do Hemmerling (2003), Myrvold (1997), and Ko (1991).

For decidability purposes, the requirement of being open or closed seems irrelevant. As noted in Section 2.4., no effective procedure can correctly decide membership in a set in a neighborhood of a boundary point. This is just as true of open and closed sets as of any other set. Open and closed sets may be of special interest, but intuitively they are no more decidable than others. Thus, if we are only concerned with decidability, strong recursiveness and TTE recursiveness carry some unnecessary baggage.

Hemmerling (2003) generalizes strong recursiveness to arbitrary sets in a natural way, using the symmetric distance function Δ_A defined above.[25] He says a set A is *Δ-decidable* if Δ_A is computable (Definition 3.4.1). As we have noted, computability of the symmetric distance function implies a kind of near-decidability. If A is strongly recursive or merely Δ-decidable, then there is an algorithm to decide whether $x \in A$ for any point $x \in \mathbb{R}^n$ except those on the boundary of A.

Yet, for decidability purposes this is still unnecessarily strong. It is just the latter implication, the existence of a decision procedure that fails only on the boundary, that interests us here. We might instead take this implication itself as a decidability concept. Myrvold calls a set *decidable ignoring boundaries* (*d.i.b.*) if its interior and exterior are both r.e. open (Definition 3.4.5). This is equivalent to the existence of a Turing machine M that outputs 1 if its input x is in A and 0 if x is not in A, unless x lies on the boundary of A, in which case M does not halt at all. It could also be defined as the existence of a machine that computes *weak characteristic function*[26]

---

[25] Hertling (1999) had studied the computability of this function but only with respect to open and closed sets.

[26] This was defined by Kummer and Schäefer (1995) but for rational arguments only, applied by them only to closed convex sets, and by Ziegler (1999) to regular sets.

$$\omega_A(x) = \begin{cases} 1 & \text{if } x \in \operatorname{int}(A), \\ 0 & \text{if } x \in \operatorname{ext}(A), \\ \text{undefined} & \text{otherwise,} \end{cases}$$

everywhere on the domain of that function and does not halt elsewhere. This turns out to be strictly weaker than Δ-decidability (Corollary 3.4.7, Proposition 3.4.8).

Ignoring boundaries seems a very appropriate cure for the emptiness of naïve decidability. That emptiness arises from the fact that we can never correctly decide membership in a set both on the boundary and near it, by the TUP. Therefore, ignoring the boundary seems to be the minimal sacrifice. Furthermore, as Myrvold writes, "There are cases in analysis and physics in which the boundaries of a given region are of little concern" (1997).

However, we can perhaps be more specific. In which cases is the boundary of little concern? They would seem to be cases where the boundary is small (though this might not always be *sufficient* reason to ignore a boundary[27]). It is easy to forget that the

[27] Myrvold himself mentions a case from quantum mechanics where the boundary of a set seems important, though it is in a sense very small (1997). The set **E** of entagled vectors in the product space of two Hilbert spaces is r.e. open (with respect to an appropriate coding) and its complement **F** is nowhere dense (in the natural topology). Hence **F** is identical to its own boundary and is d.i.b. Myrvold remarks, "This is perhaps not very interesting, since ignoring the boundary is ignoring the whole of **F**." Here the boundary is small in the topological sense of being nowhere dense (and therefore of first Baire category). It may also be small in the relevant measure, but what measure that would be is not obvious. The point is, it seems important regardless of its size, because it *is* one of the sets in question. On the other hand, its importance really depends on our motivations. In this case, since the boundary of **E** is very small in a topological sense, d.i.b. means that we can decide cases that in that same sense constitute the vast majority.

boundary of a set—those points where the symmetric distance function takes the value zero—can take up a great deal of space. The boundary of the set of rational numbers, for example, is the entire real line. Perhaps, then, we should distinguish as especially decidable those sets that are d.i.b. *and* have very small boundaries. These might be the d.i.b. sets whose boundaries are small in some appropriate measure, or if one prefers a purely topological concept, those with boundaries that are of first Baire category (countable unions of nowhere dense sets).

On the other hand, after dispensing with the tacit assumption that boundaries are small or usually small, it is not clear why they might be any less important than other sets. Penrose even gives them special importance, suggesting that a good notion of recursive set should depend in a significant way on the boundary points (1989). Granted, we are almost forced to ignore them, since we cannot correctly decide all the points in a neighborhood of a boundary point. Yet, if it was the assumption that boundaries are small that gave us license to ignore them, it seems better to base a notion of decidability on the tacit assumption itself—that is, to define a relaxed decidability in terms of the *size* of the set where an algorithm does not produce the desired output. The following sections are concerned with such notions.

2.7. Recursive approximability

For a non-trivial set of real numbers, there is no decision algorithm that always works. Recognizing this, Ko (1991) defines the notion of a recursively approximable set, a set for which there is a decision algorithm that usually or probably works. This amounts to the existence of an approximately accurate decision algorithm, in the form of a Turing machine that halts on every name for a real point and such that the set of points where the machine gives the wrong output (the *error set*) can be made arbitrarily small in Lebesgue measure (or Lebesgue outer measure).[28] The bound on the measure of the error set is specified as an input on the machine's work tape. In sum, a set is *recursively approximable* (*r.a.*) if there is a machine that, given a parameter n, (1) halts on every real input and (2) computes the set's characteristic function correctly, except perhaps on some set with Lebesgue outer measure less than 2^{-n} (Definition 3.5.1).

Ko explains that such a machine would decide a set "with an error probability less than or equal to 2^{-n}, where the probability is measured by the natural Lebesgue measure..." (1991). However, this motivating remark raises a question: Is it legitimate to equate probability with Lebesgue measure? To some extent this seems natural. If we choose a point in $\mathbb{R}^n$ at random, or if one is chosen for us by some physical system, the probability of getting a point in a particular set that has very small measure would seem to be very small. This will be discussed further in section 2.9. For now, let us only note that recursive approximability is an adaptable notion, easily relativized to any measure. (See Definition 3.5.1.) Whatever the appropriate probability measure may be, the notion of

---

[28] Lebesgue measure is the standard notion of volume in $\mathbb{R}^n$, and Lebesgue outer measure is related notion that applies to more sets. The precise definitions are not essential to the present discussion.

recursive approximability relative to that measure expresses the existence of a decision algorithm that will probably work, and further, with any success rate requested.

In addition, recursive approximability suggests a way of graphing sets in $\mathbb{R}^2$ with arbitrary accuracy, not distance-wise, as with recursive open and closed sets, but measure-wise. Consider a recursively approximable subset A of the unit square, and again let us divide the unit square into smaller squares containing pixels, as in Section 2.5. By choosing the screen resolution, i.e., the number of pixels, sufficiently high, we can effectively graph A in such a way that the *area* of the screen that is incorrectly colored is as small as we like. This will be helpful in understanding the results of Sommerer and Ott described in Chapter 4.

We will just sketch the argument. If A is r.a., there is a machine M that correctly decides A up to measure 2^{-n} for any given n. We fix n small and, by the TUP, we enumerate the open sets W_i ('W' for white) of points x on which M outputs 0 (when given n and some name ϕ for x). We also enumerate the open sets B_i ('B' for black) where M outputs 1 (given the same n and some name ϕ for a point). Finitely many of these sets $W_{i \leq j}$ and $B_{i \leq k}$ will suffice to cover the closed unit square, so we determine such a covering. Now, to accurately graph A, we need only pixelate the unit square with fine enough resolution to very accurately graph the sets $W_{i \leq j}$ and $B_{i \leq k}$. Where the $W_{i \leq j}$ and $B_{i \leq k}$ overlap, the coloring of pixels may be arbitrary, for these regions will be very small. Thus, in $\mathbb{R}^2$, r.a. implies the existence of an effective method to correctly plot a set up to arbitrarily small area.

As we saw in Section 2.5, the notion of a recursive closed set also implies an effective method for plotting sets in $\mathbb{R}^2$ with arbitrary accuracy, but in a different sense. The latter implies that a set in $\mathbb{R}^2$ can be plotted up to arbitrarily small *distance*, which is not equivalent to area-wise accuracy. The proof of Proposition 3.6.6 exhibits a set that is recursive closed and even strongly recursive but not r.a., namely a generalized Cantor set with measure equal to a non-computable number. This set cannot be effectively plotted up to arbitrarily small area, but it can be plotted up to arbitrarily small distance. Likewise, there are sets that can be plotted with arbitrary accuracy measure-wise, but not distance wise. Take any non-recursive set K of natural numbers, and consider the set $A = \{(x, 2^{-k}) \in \mathbb{R}^2\colon\ x \in [0, 1], k \in K\}$. This closed set is not recursive, but since it has measure zero, it is graphed with accuracy up to measure zero by a completely blank screen.

Which sort of accuracy is most useful just depends on one's intentions. The sets that most concern this dissertation are sets of points that represent physical states, and ultimately we would like to apply our knowledge of such sets to actual systems encountered in the real world. If we assume that the states of actual systems that we encounter will be randomly distributed over a continuum, then a measure-wise accurate picture of a set A seems more immediately useful than a distance-wise accurate one, for the former gives us a sense of how likely it is that the state of an arbitrary system will lie in A, and further, how likely it is that a state *in a given small region* of state space will lie in A. It enables us to classify states with a high *probability* of accuracy (assuming we have chosen an appropriate measure).

2.8. Decidability up to measure zero

Recursive approximability is clearly a valuable property, but as an analog of classical decidability it is somewhat unsatisfying due to its merely approximative nature. In contrast, the classical concept of decidability in discrete recursion theory (or for that matter, logic, as in Gödel [1931] 1986, [1934] 1965) is absolute and total. It concerns what can be decided in *all* cases using a single algorithm (or effectively axiomatized theory). Such complete decidability is not possible for non-trivial sets of real numbers, so Ko settles for an arbitrarily small but non-zero probability of error. But must we be satisfied with this? It might yet be possible do decide non-trivial properties with *exactly zero* probability of error, since probability zero is not the same as impossibility. This would be intuitively better than a small non-zero chance of error and closer in spirit to the absolute ideal of classical decidability.

There is an obstacle to such thorough decidability built into the concept of r.a. If we require a decision procedure that always halts, as Ko does, only sets of trivial measure (either zero or the same as the full background space) can be decided all the way up to measure zero (Parker 2003). If a decision procedure for a set A halts on every input representing a real point, then by the TUP, it must make mistakes near the boundary of A. Furthermore, if there is a boundary point where every neighborhood contains positive-measure portions of both A and A^C, then the algorithm must make mistakes on a positive-measure set of points. But the only sets in $\mathbb{R}^n$ that do not have such boundary points are measure-theoretically trivial, i.e., sets A such that either A or A^C has measure zero.[29]

The motivations for Ko's requirement that a decision algorithm must halt on every point are not apparent. Assuming we have a machine that could possibly give incorrect output anyway, the epistemological situation would seem no worse if in principle that machine could also fail to halt, but only with probability zero. This would not affect the probability of obtaining a correct output, and in application we could confidently assume that non-halting cases would never arise.

Therefore, why not simply demand a machine that, with probability one (whatever the probability measure may be), will halt and give correct output? Let us say a set is *decidable up to μ-measure zero* (or μ-d.m.z), if some Turing machine will compute its characteristic function except on a (possibly empty) *error set* with μ-measure zero, where the machine might decide incorrectly or not at all (Definition 3.5.4). We will use λ for Lebesgue measure and therefore write 'λ-d.m.z.' for decidability up to Lebesgue measure zero, or if the measure is clear from the context, we may just write 'd.m.z.'

Our intent in defining μ-d.m.z. is that μ should be chosen to reflect some

---

[29] For these sets, decidability up to measure zero is obvious anyway. A measure-zero set is correctly decided, up to measure zero, by the Nancy Reagan algorithm: Just say no.

probability that interests us. We will say more below about what probability we have in mind. Suffice it to say for the moment that we would *like* it to be the case that if a set A of states for some class of physical systems has measure zero, then the probability that the state of any particular system we come across will lie in that set is zero. In that case, d.m.z. implies that an algorithm will succeed with probability one in any such application. For this purpose, our chosen measure μ need not be exactly equal to the probability that concerns us, but it should be chosen so that the probability is *absolutely continuous* in μ—all sets with μ-measure zero should be assigned probability zero. Notice, we are not *asserting* a connection between "phase space averages" and "time averages" for a particular system, nor any other relation between one measure or probability to another. We are merely stating a *preference* that μ should reflect the very probability that interests us, namely the probability that the state of any one system we encounter will fall in a given set.

We will see in Chapter 3 that λ-d.m.z. is strictly stronger than r.a.; it implies r.a., but the converse fails (Theorem 3.5.10 and Proposition 3.6.3). One may find this surprising, since r.a. requires an algorithm that halts on *every* point, while d.m.z. does not. However, given an algorithm that halts on *almost* every point and correctly decides a set A on almost every point, it turns out that we can effectively construct an arbitrarily small r.e. open set that covers those points where the algorithm does not halt. We then assign arbitrary output to the points in that r.e. open set. Combining this with the original algorithm yields an algorithm that halts everywhere and errs only on a set of arbitrarily small measure.

Note also that since λ-d.m.z. implies r.a., it also implies, for sets in $\mathbb{R}$ or $\mathbb{R}^2$ at least, an effective method of graphing up to arbitrary accuracy, measure-wise. (It does not imply that a set can be graphed with measure-*zero* error, at least not with finitely many pixels of the same size and shape.)

It is not claimed here that d.m.z. is the best concept of decidability in every respect for every real-valued context. In fact, r.a. is perhaps more pragmatically motivated, for an extremely minuscule probability of error is usually good enough for practical purposes. However, it is just this pragmatism that makes r.a. less analogous to classical decidability than d.m.z. Though only trivial sets of reals are decidable in the absolute, naïve sense, clearly those sets that are decidable all the way up to measure zero come closer to that standard than those that can only be decided up to an arbitrarily small non-zero measure.

Also, because d.m.z. is stronger than r.a., it distinguishes a higher level of decidability to which r.a. is insensitive. The proof of Proposition 3.6.3 exhibits a set that is r.a. but not d.m.z., and in Chapters 4 and 5 we will see mathematical models of physical systems such that sets of states corresponding to certain qualitative properties are r.a. but not d.m.z. The fact that these systems are not d.m.z. bears out intuitions that their behavior is undecidable, intuitions not captured by the concept of r.a.

One can define still stronger notions of decidability that might be preferable in some respects. Strictly speaking, though, no stronger decidability could entail any greater *probability* of correct output. D.m.z., with the right measure, guarantees an algorithm that succeeds with probability one, the best *rate* of success we can hope for. Furthermore, a stronger notion of decidability would imply a weaker notion of *un*decidability, and this would be undesirable for our purposes. One of our tasks in Chapters 4 and 5 is to find the most meaningful *un*decidability results possible for certain dynamical systems.

Perhaps an extreme purist would not much like d.m.z., nor r.a. Both concepts ride roughshod over some of the subtleties we have discussed. They have the right sorts of symmetry, and intuitively simple sets come out d.m.z. and r.a., but also, *any* measure-zero set is both d.m.z. and r.a., no matter how complex it may be. Yet if what we want to know is whether we can apply a single decision procedure to various systems that we might encounter in the world, the exact states of which we cannot know in advance, probabilistic or measure-theoretic notions of decidability seem appropriate.

2.9. Probability and measure

We noted in 2.7 that the motivation for recursive approximability (with respect to Lebesgue measure) depended on the assumption that Lebesgue measure is somehow related to probability. In physical applications, we would like to suppose that at any given time, a given physical quantity is unlikely to take on a value in any particular set that has small Lebesgue measure. To motivate λ-d.m.z., we made a related claim: that barring "special circumstances," a quantity *will not* be in a given set with *zero* Lebesgue measure at any given time. What right do we have to make such claims? What does

Lebesgue measure have to do with probabilities?

Before even considering such questions, we should clarify what sort of probability concerns us—the probability of what? Suppose we have some model of a physical system in which the possible states of the system are represented as points in $\mathbb{R}^n$. Suppose also that we have an algorithm that partly decides the membership of a certain set of states in that model. We would like to be confident that whenever we encounter one of the systems for which our model is intended, it will be in one of those states for which our decision procedure works. Take for example a system of three bodies in otherwise empty space, evolving by gravitational attraction. Idealizing the bodies as point masses, we can represent the state of such a system in $\mathbb{R}^{18}$, allotting three dimensions for the position of each body and three for each velocity. For the sake of comparison among different three-body systems, let us also include the three masses in our state space, so that each state is a point in $\mathbb{R}^{21}$. Now suppose we have a procedure to decide, say, whether a body will ever escape such a system, but there is some small set of states on which the procedure does not work. Then we would *like* to know that whenever we encounter a three-body system in the real world, chances are good that its state will not be in that set. So the probability that mainly concerns us is the probability that, *when* we stumble on one of the systems for which a certain model is intended, its state will lie in a certain set.

We might also be concerned with systems that are deliberately prepared for the sake of an experiment. In that case we can usually assume some error in the preparation. If we again have an algorithm to determine something about the behavior of the prepared system, but this algorithm fails on some small set of states, then we would *like* to know that, for an arbitrary preparation attempt, the chances of preparing a state in the error set is very small.

The assumption that sets of measure zero, in Lebesgue measure and other related measures, should be assigned probability zero is often made in statistical mechanics, though its justification is a persistent problem in philosophy of science (Sklar 1973, 1993; Malament and Zabell 1980; von Plato 1983; Batterman 1998; Vranas 1998). Sklar devotes a section of his book *Physics and Chance* (1993, 182-188) to it but offers no definitive solution and deems it an interesting outstanding problem (414). A common explanation for the success of equilibrium statistical mechanics (specifically, Gibbs phase averaging) assumes that one can ignore sets with zero measure in the so called micro-canonical measure on an energy surface. These are the same sets that have zero Lebesgue measure.[30] That account of statistical mechanics is attacked by Malament and Zabell (1980), but their own account also claims that the relevant probability measure on a state space will be absolutely continuous. They write that absolute continuity "seems to be

[30] Lebesgue measure is not well-defined on an arbitrary manifold, but it is well-defined relative to a coordinate system on an open subset of a manifold. If the manifold is C^1 then in any two coordinate systems on the same neighborhood, the same sets will have zero Lebesgue measure (Folland 1984, 332).

universally accepted as a basic postulate of statistical mechanics..." (345). The success of statistical mechanics and the role of absolute continuity in the leading explanations of that success give the postulate some support.

Malament and Zabell go on to offer some justification for absolute continuity in terms of the way in which a system can be prepared. They show that in $\mathbb{R}^n$ at least, it is equivalent to a property called *displacement continuity*: that the probability associated with a set be very close to that of another set obtained by a small translation. This property is plausible, they suggest, because of our inability to prepare a system with perfect accuracy (and because, when we *can* prepare a system so as to lie exactly on a particular manifold, we can then reduce the state space to just that manifold with the restricted topology and measure). This preparation argument would seem to be supported by the Gaussian theory of errors. Assuming that the errors in preparation (or measurement) are many, small, independent, and non-systematic (i.e., not inclined to one direction more than another), we obtain a Gaussian normal distribution (Poincaré [1896] 1912, 128 ff; [1907] 1953, 77), which is indeed displacement continuous and absolutely continuous in Lebesgue measure.

Such considerations speak to the tenability, popularity, and usefulness of associating measure zero with probability zero. Ultimately, though, they are inconclusive. For all we know, the states of physical systems in this world might be confined to some discrete set, such as the rational points in some coordinate system, and our continuous models might only approximate such systems. In that case, the most accurate probability measure would assign a positive probability to the set of rational points, a set of zero Lebesgue measure. Thus the assumption of absolutely continuous probabilities is a hypothesis of roughly the same status as the continuity of state spaces. It is an attractive hypothesis worthy of special consideration, and it forms part of the background of this investigation. If, in the end, we *can* reasonably ignore sets of small measure, as many have supposed, then r.a. and d.m.z. are useful notions.

Again, though, both r.a. and d.m.z. can be relativized to any measure. Whatever the most appropriate probability measure for a given state space may be, these concepts of decidability can be adapted accordingly. Specific results concerning the decidability or undecidability of particular sets with respect to Lebesgue measure, such as those presented in Chapters 3 through 5, might not survive such modulations, but the value of the *general* concepts is not threatened by any doubts about particular probability measures. If a set is d.m.z. in the most appropriate probability measure, or even one that merely assigns measure zero to the right sets, then there is a decision algorithm that one can trust.

2.10. Conclusions

Many different notions of recursive and decidable sets of real numbers are possible, and each has its virtues and vices.

Here we have mainly considered notions founded on Turing's discrete conception of computation, i.e., the manipulation of finite symbol strings, as modeled by the Turing machine. Though we did not consider all possible concepts of computation, the real RAM model was found not to be a good model for the way we usually carry out computations, whether by hand or by machine. In particular, real RAMs can carry out exact comparisons between real numbers, which in practice we cannot generally do, even with the aid of machines.

Within the discrete conception of computation, we first considered the naïve notion of a decidable set of reals, which consists in the existence of a foolproof decision procedure. It would not be unreasonable to say that this is *the* natural notion of decidability over the reals, but it is of little use, since only the most trivial subsets of $\mathbb{R}^n$ are decidable in this sense.

The TTE notion of a recursive open or closed set is more forgiving while still implying some useful computability properties and maintaining some analogies with the notion of a recursive set of natural numbers. However it does not imply even an approximate decidability, but only a semi-decidability analogous to recursive enumerability. Strong recursiveness recovers the symmetry of decidability with respect to complementation but is perhaps too strong for a notion of decidability. Its attractive decidability implication is just decidability ignoring boundaries, but it is strictly stronger than this notion and applies only to open and closed sets.

The notion of Δ-decidability, or the computability of the symmetric distance function, generalizes strong recursiveness to arbitrary sets in $\mathbb{R}^n$, but is still stronger than the implication that makes it attractive as a notion of decidability, namely decidability ignoring boundaries. The latter notion cuts more to the heart of the matter and seems minimally permissive, for it is at the boundary points of a set that correct decisions are always impossible. Yet we must remember that boundaries can be large regions, and it is often unsatisfactory to ignore them.

In physical applications it seems we are more concerned to have a decision procedure that only fails on a small set of points, on the assumption that such a set will be associated with a small probability. Recursive approximability, a sort of decidability up to arbitrarily small measure, is therefore very practical. However, it involves the somewhat arbitrary requirement that a decision procedure should always halt, right or wrong, and it does not minimize errors in the strongest way possible. Decidability up to measure zero is a stronger notion and closer in spirit to the classical notion of decidability, but not so strong as to be uninteresting. If we choose a measure μ in which the probability distribution over a set of physical states is absolutely continuous, then μ-d.m.z. implies that some decision procedure succeeds with probability one. For practical purposes, this seems to be the closest thing possible to decidability.

In Chapters 4 and 5 we will consider applications of these notions, especially d.m.z., to systems that have been thought to have intuitively undecidable behavior. We will see in both cases that it is d.m.z., or rather the lack thereof, that expresses the intuitive undecidability that those systems exhibit.

CHAPTER 3

THE LOGICAL RELATIONS BETWEEN VARIOUS NOTIONS OF A COMPUTABLE SET

3.1. Introduction

In this chapter we establish the relations of implication, or more often the lack thereof, among various notions of a decidable or recursive set of real numbers. The main point is that the notion of decidability up to measure zero is not equivalent to other notions of decidability or recursiveness, such as recursive approximability, decidability ignoring boundaries, or "recursive closed." The logical relations that do hold among such concepts are summarized in Figure 3.3 at the end of the chapter. Though it is conceptually self-contained, this chapter may be regarded as a technical appendix to Chapter 2 (or Chapter 2 as a motivating introduction to this one).

Our approach to recursive analysis is similar in spirit, though not in every detail, to the so-called Type-2 Theory of Effectivity (TTE), of which Weihrauch is a leading proponent. Following Weihrauch, we generalize notions of decidability or recursiveness to sufficiently structured topological spaces (second-countable T_0-spaces), discussing examples in Euclidean space as special cases. Most of the definitions and results of Sections 2 and 3 are based on those of Weihrauch (2000), along with some of Ko's

(1991) notation and concepts.[1] Several notions are slightly modified in hopes of greater simplicity. The proofs of many propositions throughout this chapter are given elsewhere and only references are given here. Results not otherwise attributed, to the best of the author's knowledge, have not been proven elsewhere. Certainly all results involving d.m.z. are original.

Before beginning, let us institute some basic notation. We use the backslash '\' to denote set subtraction, the bold Roman superscript '$^{\mathbf{C}}$' for complementation, and $\mathbf{P}(A)$ for the power set of A, i.e., the set of all subsets of A. A^B denotes the set of functions $f: B \to A$. A function $f: \subseteq A \to B$ is one defined on some subset of A. We use dom(f) for domain, ran(f) for range, and $f^{-1}(x)$ for the set $\{y \in \text{dom}(f): f(y) = x\}$. A function $\phi: \mathbb{N} \to A$ may be written as a sequence $\{\phi_i\} = (\phi_0, \phi_1, \ldots)$, where each ϕ_i denotes $\phi(i)$, and we may write $a \in \phi$ to mean $a \in \text{ran}(\phi)$, and $\phi \subseteq A$ for ran(ϕ) $\subseteq A$.

We will also need a very little measure theory. Intuitively, a measure is a notion of volume, or quantity of space. An *outer measure* on a set X is a function $\mu: \mathbf{P}(X) \to [0, \infty)$ such that $\mu\varnothing = 0$, and for any sequence $\{B_i\} \subseteq \mathbf{P}(X)$, $B_i \subseteq B_j \Rightarrow \mu B_i \leq \mu B_j$, and $\mu \bigcup_{i=0}^{\infty} B_i \leq \sum_{i=1}^{\infty} \mu B_i$. A *measure* is a function $\mu: \subseteq \mathbf{P}(X) \to [0, \infty)$, with the same properties as an outer measure and a little more: the final inequality becomes an equality. However, the domain of a measure is not generally $\mathbf{P}(X)$ but some *sigma-algebra*, a subset of $\mathbf{P}(X)$ containing $\varnothing$ and X and closed under complementation and countable

[1] As noted in Section 2.3.1, Weihrauch's and Ko's contributions go back at least to 1981. Here we rely on their more recent books, Weihrauch 2000 and Ko 1991, where earlier references can be found.

union. We will reserve λ for Lebesgue measure, the standard measure in $\mathbb{R}^n$. It is equal to ordinary length in $\mathbb{R}$, area in $\mathbb{R}^2$, and n-dimensional volume in $\mathbb{R}^n$ for all cases where these intuitive notions are clear. We omit the precise definition, which is inessential to our discussion and can be found in any analysis text.

Finally, a bit of topology: A *topological space* is a pair (X, τ) where X is a set and τ a *topology*, i.e., a set $\tau \subseteq \mathbf{P}(X)$ closed under union and finite intersection and containing as elements $\varnothing$ and X. The elements of X are called *points,* the elements of τ are called *open sets,* and an open set containing a point x is called a *neighborhood* of x. A *closed set* is the complement $U^{\mathrm{C}} = X \setminus U$ of an open set U. The *interior* $\mathrm{int}(A)$ of $A \subseteq X$ is the largest open subset of A, the *exterior* $\mathrm{ext}(A)$ is the largest open subset of $A^{\mathrm{C}} = X \setminus A$, and the *closure* $\mathrm{cl}(A)$ is the smallest closed set containing A. The *boundary* ∂A of A is $\mathrm{cl}(A) \setminus \mathrm{int}(A)$. A set $\beta \subseteq \tau$ is a *base* for τ if every open set (element of τ) is a union of members of β. A set $\sigma \subseteq \tau$ is a *subbase* for τ if τ is the smallest topology containing σ, i.e., if τ is *generated* from σ just by closing σ under union and finite intersection. A topological space is *second-countable* if it has a countable base, and finally, a T_0*-space* is a topological space (X, τ) such that for each $x, y \in X$,

$$\{U \in \tau:\ x \in U\} = \{U \in \tau:\ y \in U\} \iff x = y.$$

3.2. Naming systems and coded topological spaces

3.2.1. Codings and spaces

We will represent an element of a countable set by a natural number. To fix some concrete codings, let $\langle a,b\rangle = (a^2 + 2ab + b^2 + 3a + b)/2$, for all $a,\ b \in \mathbb{N}$. This is a recursive one-to-one mapping of $\mathbb{N} \times \mathbb{N}$ onto $\mathbb{N}$, and its inverse is also recursive. Let $\langle a_1, a_2, \ldots, a_n\rangle = \langle a_1, \langle a_2, \ldots, \langle a_{n-1}, a_n\rangle\rangle \ldots\rangle$. Now for all $a,\ b \in \mathbb{N}$, let

$$z_{\langle a,b\rangle} = \begin{cases} b & \text{if } a = 0, \\ -b & \text{otherwise,} \end{cases}$$

and

$$q_{\langle a,b\rangle} = \begin{cases} 0 & \text{if } b = 0, \\ \dfrac{z_a}{b} & \text{otherwise.} \end{cases}$$

Thus $\{z_i\}$ is an effective enumeration of the integers, and $\{q_i\}$ of the rationals.

In the context of an uncountable space, where there is no way to represent each point with a distinct whole number, we represent a point by a sequence of approximations. This requires some minimal notion of nearness, i.e., a topology, and a way of picking out points with countable sequences. Therefore we follow Weihrauch in taking as our basic domains second-countable T_0-spaces, i.e., topological spaces such that every point is uniquely picked out by a countable sequence of open sets containing it. (See Section 3.1.)

Definition 3.2.1 (see Weihrauch 2000).[2] A *coded topological space* is a pair $\mathbf{X} = (X, \sigma)$ where

(i) X is a non-empty set,
(ii) $\sigma = \{\sigma_i\}_{i \in \mathbb{N}}$ is a sequence of subsets of X, and
(ii) $\{i\colon x \in \sigma_i\} = \{i\colon y \in \sigma_i\} \iff x = y$.

Note that σ here is not necessarily a topology, but a countable subbase for a second-countable T_0-space $(X, \tau_{\mathbf{X}})$. (See Section 3.1.) Throughout this chapter, X refers to the domain of a coded topological space $\mathbf{X} = (X, \sigma)$, and most definitions are implicitly relativized to a particular enumeration $\{\sigma_i\}$ and the topology $\tau_{\mathbf{X}}$ generated by it.

To treat the set of real numbers as a coded topological space, we need a countable subbase. Let

$$I_{\langle a,b\rangle} = \begin{cases} \text{the interval } (q_a, q_b) & \text{if } q_a < q_b, \\ \varnothing & \text{otherwise.} \end{cases}$$

Thus $I = \{I_i\}$ is an effective enumeration of the open intervals (q_j, q_k) with rational endpoints. To extend this to $\mathbb{R}^n$, define an enumeration $I^n = \{I_i^n\}$ of the *open rational n-intervals* $(q_{a_1}, q_{b_1}) \times (q_{a_2}, q_{b_2}) \times \ldots \times (q_{a_n}, q_{b_n})$ by

$$I^n_{\langle a_1,b_1,a_2,b_2,\ldots,a_n,b_n\rangle} = \begin{cases} (q_{a_1}, q_{b_1}) \times (q_{a_2}, q_{b_2}) \times \ldots \times (q_{a_n}, q_{b_n}) & \text{if for all } j \le n,\ q_{a_j} < q_{b_j}, \\ \varnothing & \text{otherwise.} \end{cases}$$

[2] This is equivalent to Weihrauch's notion of an "effective" topological space, but that term is perhaps misleading, since neither the enumeration σ nor its range is required to be effective in any sense. When useful effectiveness conditions are imposed, one obtains a *computable* topological space (Weihrauch 2000).

Wherever we discuss $\mathbb{R}^n$ below, we have in mind the coded topological space ($\mathbb{R}^n$, I^n) and the usual Euclidean topology $\tau_{\mathbb{R}^n}$, which is generated by I^n.

A point in a coded topological space can be represented by an infinite list of its neighborhoods, or more precisely, by a sequence $\phi \in \mathbb{N}^{\mathbb{N}}$ that lists the indices of all subbase elements containing the point. Hence,

Definition 3.2.2 (Weihrauch 2000). (i) A *representation* of a set X is a function $\rho: \subseteq \mathbb{N}^{\mathbb{N}} \to X$ onto X.

(ii) The *standard representation* $\rho_{\mathbf{X}}: \subseteq \mathbb{N}^{\mathbb{N}} \to X$ of a coded topological space $\mathbf{X} = (X, \sigma)$ is defined by $\rho_{\mathbf{X}}(\phi) = x \Leftrightarrow (i \in \phi \Leftrightarrow x \in \sigma_i)$. If $\rho_{\mathbf{X}}(\phi) = x$, we call ϕ a *standard name* for x.

(iii) A *naming system* for a set A is either an enumeration of A or a representation of A.[3] For any naming system α, ϕ is an *α-name* for x if $\alpha(\phi) = x$.

For the real numbers in particular, the standard representation $\rho_{\mathbb{R}}$ is one of several natural representations:

Definition 3.2.3 (Representations of $\mathbb{R}$; see Weihrauch 2000).

(i) (Standard) $\rho_{\mathbb{R}}(\phi) = x \in \mathbb{R} \Leftrightarrow (i \in \phi \Leftrightarrow x \in I_i)$. Similarly, $\rho_{\mathbb{R}^n}(\phi) = x \in \mathbb{R}^n$

[3] For Weihrauch (2000), a naming system is a *notation* or representation, the former being a function on a set of finite symbol strings. Here we substitute numbers for finite strings, so notations can be replaced with ordinary enumerations.

$\Leftrightarrow (i \in \phi \Leftrightarrow x \in I_i^n)$.

(ii) (Base-b digital expansion) Assume $\phi = (a_1\, a_2 \ldots a_m \,.\, a_{m+1}\, a_{m+2} \ldots)$, where $a_1 \in \{-, 1, 2,\ldots, b-1\}$, and for $i > 1$, $a_i \in \{0,\ldots, b-1\}$. (Technically, then, ϕ is not an element of $\mathbb{N}^{\mathbb{N}}$, but we may so regard it by identifying the decimal point with the number b and the negative sign with $b + 1$.) Then let

$$\rho_{\text{base-}b}(\phi) = \begin{cases} -\sum\limits_{i=2}^{\infty} a_i b^{m-i} & \text{if } a_1 = -, \\ \sum\limits_{i=1}^{\infty} a_i b^{m-i} & \text{otherwise.} \end{cases}$$

(iii) (Inclusive regular Cauchy) Let $\rho_C(\phi) = x \Leftrightarrow (\forall i \in \mathbb{N})\, |q_{\phi_i} - x| \le 2^{-i}$. We call $\phi \in \rho_C^{-1}(x)$ a *regular Cauchy function*[4] or *regular Cauchy sequence* for x.

(iv) (Strict regular Cauchy) Let $\rho_C'(\phi) = x \Leftrightarrow (\forall i \in \mathbb{N})\, |q_{\phi_i} - x| < 2^{-i}$. We call $\phi \in \rho_C'^{-1}(x)$ a *strict regular Cauchy function* or *strict regular Cauchy sequence* for x.

We will consider the relative merits of these representations in Section 3.2.3.

---

[4] Weihrauch and Ko simply call these *Cauchy* functions (sequences, representations), without the qualifier 'regular.' However, the requirement of uniformly rapid convergence (here, the bound 2^{-i}) is not part of the standard notion of a Cauchy function, and is similar to Bishop's notion of regularity (1967, 15). Constructivists such as Aberth, and Bishop himself (p. 85), also tend to use the phrase 'Cauchy sequence' with the tacit implication that the rate of convergence is ruled by some effective function of the indices, due to constructivist background principles.

3.2.2. Machines and computable functions

Our computability and decidability concepts will be based on the notion of a computer program or machine supplied with an infinite input sequence, and perhaps also a distinguished finite input or parameter. We use the following notation for functions computed by machines:

Definition 3.2.4 (cf. Ko 1991). Let $\phi \in \mathbb{N}$ or $\mathbb{N}^{\mathbb{N}}$ and $a, b \in \mathbb{N}$. Then,

(i) If M is a machine that, given input ϕ, outputs a and halts, we write $M(\phi)\downarrow$ and $M(\phi) = a$.[5] If M does not halt given input ϕ, we write $M(\phi)\uparrow$.

(ii) If, given inputs ϕ and a, M outputs b and halts, we write $M(\phi, a)\downarrow$ and $M(\phi, a) = b$. If M does not halt given inputs ϕ and a, we write $M(\phi, a)\uparrow$.

(iii) If $M(\phi, i)\downarrow$ for all $i \in \mathbb{N}$, we write M^{ϕ} for the sequence of outputs $\{M(\phi, i)\}_{i \in \mathbb{N}}$.

Ko (1991) and Weihrauch (2000) define infinite-input machines in detail. Here we instead rely on the generalized Church-Turing Thesis stated in Section 2.3.1. To repeat the thesis more formally,

> **Generalized Church-Turing Thesis for infinite arguments:** The functions $M: \subseteq(\mathbb{N}^{\mathbb{N}} \times \mathbb{N}) \rightarrow \mathbb{N}$ computed by an infinite-input machine are precisely those that can be computed by any effective procedure.

[5] Note that $M(\phi) = a$ implies $M(\phi)\downarrow$. In cases where $M(\phi)$ is not defined, we will regard $M(\phi) = a$ as a false statement, rather than a category error lacking any truth value. Thus we can write, for example, "If $M(\phi) = a$ then..." without qualms.

By appeal to this assumption, we ignore the details of machines and give computability arguments by describing effective procedures in intuitive language.

Now we define computability for a few types of functions. (See Figure 3.1.)

Definition 3.2.5 (see Weihrauch 2000). (i) Let α be a naming system (an enumeration or representation) for some set A and $\nu = \{\nu_i\}$ an enumeration of a countable set B. A function $f: \subseteq A \to B$ is (α, ν)-*computable* if there is a machine M such that for all $\phi \in \mathrm{dom}(f \circ \alpha)$, $\nu \circ M(\phi) = f \circ \alpha(\phi)$. In that case M is said to (α, ν)-*compute* f.

(ii) Let α be a naming system for some set A and ρ a representation for B. A function $f: \subseteq A \to B$ is (α, ρ)-*computable* if there is a machine M such that for all $\phi \in \mathrm{dom}(f \circ \alpha)$, $\rho(M^\phi) = f \circ \alpha(\phi)$, in which case M is said to (α, ρ)-*compute* f.

(iii) Let $\mathbf{X} = (X, \sigma^{\mathrm{X}})$, $\mathbf{Y} = (X, \sigma^{\mathrm{Y}})$ be coded topological spaces with standard representations ρ_{X} and ρ_{Y}. A function $f: \subseteq X \to Y$ is *computable* if it is $(\rho_{\mathrm{X}}, \rho_{\mathrm{Y}})$-computable.

$$\begin{tikzcd} \subseteq A \arrow[r, "f"] & B \\ \mathbb{N} \text{ or } \mathbb{N}^{\mathbb{N}} \arrow[u, "\alpha"] \arrow[r, "M"'] & \mathbb{N} \arrow[u, "\nu"'] \end{tikzcd} \qquad \begin{tikzcd} \subseteq A \arrow[r, "f"] & B \\ \mathbb{N} \text{ or } \mathbb{N}^{\mathbb{N}} \arrow[u, "\alpha"] \arrow[r, "\phi \mapsto M^\phi"'] & \mathbb{N}^{\mathbb{N}} \arrow[u, "\rho"'] \end{tikzcd}$$

Figure 3.1. Computability of functions. If ν is an enumeration of B (left), then 'f is (α, ν)-computable' means that some machine takes each α-name for a to a ν-name for $f(a)$. If ρ is a *representation* of B (right), then 'f is (α, ρ)-computable' means there is a machine M such that if ϕ is an α-name for a then the *sequence* $M^\phi = \{M(\phi, i)\}_{i \in \mathbb{N}}$ is a ρ-name for $f(a)$.

Part (ii) of this definition says that a function f is (α, ρ)-computable if for some machine M, the sequence of outputs $\{M(x, i)\}_{i \in \mathbb{N}}$ represents $f(x)$. By part (iii), a function $f: \subseteq \mathbb{R} \to \mathbb{R}$ is computable simpliciter if it is computable in the standard representation $\rho_{\mathbb{R}}$.

In later sections, we will make use of the notion of a computable number, so we need the following definition and subsequent fact.

Definition 3.2.6 (Weihrauch; cf. Turing [1936-1937] 1965). (i) Let ρ be a representation for X. A point $x \in X$ is ρ-computable if there is a machine M such that $\rho(\{M(i)\}_{i \in \mathbb{N}}) = x$.

(ii) A number $x \in \mathbb{R}$ is *computable* if x is $\rho_{\mathbb{R}}$-computable.

Theorem 3.2.7 (Weihrauch). A number $x \in \mathbb{R}$ is computable iff x is $\rho_{\text{base-}b}$-computable.

Proof. See Weihrauch 2000, 93. ■

3.2.3. Comparison of representations

Of the representations defined above for real numbers, Weihrauch tends to favor his "standard" representation $\rho_{\mathbb{R}}$, while Ko employs the inclusive regular Cauchy representation ρ_C. The present author has elsewhere used the strict regular Cauchy representation ρ_C' (Parker 2000). However, Weihrauch has shown that all of these are equivalent in the following sense:

Definition 3.2.8 (Weihrauch). Two naming systems α, β are (computationally) *equivalent* ($\alpha \equiv \beta$) if ran(α) = ran(β) and the identity function on ran(α) is both (α, β)- and (β, α)-computable.

So equivalence of naming systems just means that each can be effectively translated into the other. As it turns out,

Lemma 3.2.9 (Weihrauch; see 2000, 88). $\rho_{\mathbb{R}} \equiv \rho_C \equiv \rho_C'$.

This implies immediately that for $\rho = \rho_C$ or ρ_C' and for any naming system α, a function is (ρ, α)-computable if and only it is ($\rho_{\mathbb{R}}$, α)-computable, and (α, ρ)-computable if and only it is (α, $\rho_{\mathbb{R}}$)-computable.

We do not adopt the familiar system of base-10 or base-2 digital expansions, in part because for any base b, some very simple functions are not ($\rho_{\text{base-}b}$, $\rho_{\text{base-}b}$)-computable. Addition and subtraction are not computable with respect to base-b representations, but to state this precisely, we would have to introduce the notion of a base-b representation of $\mathbb{R}^2$. Instead we show, as an example, that addition of the constant 2/3 is not ($\rho_{\text{base-}10}$, $\rho_{\text{base-}10}$)-computable.

Proposition 3.2.10 (see Weihrauch 2000). The function $f(x) = x + 2/3$ is not ($\rho_{\text{base-}10}$, $\rho_{\text{base-}10}$)-computable.

Proof. Suppose there is a machine M that $(\rho_{\text{base-10}}, \rho_{\text{base-10}})$-computes f. This means that if ϕ is a decimal for x then for each i, $M(\phi, i)$ is the i^{th} digit of a decimal for $x + 2/3$. If $\phi = 0.333\ldots$ then $\{M(\phi, i)\}_{i \in \mathbb{N}} = 0.999\ldots$ or $1.000\ldots$, so in particular, $M(\phi, 0) = 0$ or 1. Given input $(\phi, 0)$, the machine M scans only some finite initial segment r of ϕ before halting. (See Use Principle below.) Therefore, $M(\phi, 0) = M(r999\ldots, 0) = M(r000\ldots, 0)$. But $r000\ldots$ represents a number $y < 1/3$, so $y + 2/3 < 1$ and $M(r000\ldots, 0) = 0$, while $r999\ldots$ represents a number $z > 1/3$, so $z + 2/3 > 1$ and $M(r999\ldots, 0) \neq 0$. Therefore $M(r999\ldots, 0) \neq M(r000\ldots, 0)$, and we have a contradiction. ∎

It turns out that this does not matter much for *decidability* concepts. Intuitively, a set A is decidable if its characteristic function,

$$\chi_A(x) = \begin{cases} 1 & \text{if } x \in A \\ 0 & \text{otherwise,} \end{cases}$$

which is a function into $\{0, 1\}$, is computable. Those concepts best regarded as relaxed *decidabilities*, i.e., those equivalent to the existence of a relatively successful decision procedure and requiring no other conditions, can be similarly expressed in terms of the computability of characteristic functions over slightly restricted domains. Hence we are still mainly interested in the computability of functions into $\{0, 1\}$. It is easy to see from results of Weihrauch that for that purpose, digital expansions and the other three representations above are all equivalent. Weihrauch shows,

Theorem 3.2.11 (see Weihrauch 2000, 93). $f: \subseteq \mathbb{R} \to \mathbb{R}$ is $(\rho_{\text{base-}b}, \rho_{\mathbb{R}})$-computable iff it is $(\rho_{\mathbb{R}}, \rho_{\mathbb{R}})$-computable.

What concerns us at the moment, though, is not $(\rho_{\text{base-}b}, \rho_{\mathbb{R}})$-computability but $(\rho_{\text{base-}b}, \nu_{\{0,1\}})$-computability, for some reasonable enumeration $\nu_{\{0,1\}}: \mathbb{N} \to \{0, 1\}$. Let us say,

Definition 3.2.12. For all $i \in \mathbb{N}$, let

$$\nu_{\{0,1\}}(i) = \begin{cases} 0 & \text{if } i = 0, \\ 1 & \text{otherwise.} \end{cases}$$

Then in fact,

Proposition 3.2.13. Given any naming system α for a set A, a function $f: \subseteq A \to \{0, 1\}$ is $(\alpha, \rho_{\mathbb{R}})$-computable if and only if it is $(\alpha, \nu_{\{0,1\}})$-computable.

Proof. $\Rightarrow$: Suppose f is $(\alpha, \rho_{\mathbb{R}})$-computable. Fix a machine M that $(\alpha, \rho_{\mathbb{R}})$-computes f (i.e., for all $\phi \in \text{dom}(f \circ \alpha)$, $\rho_{\mathbb{R}}(M^{\phi}) = f \circ \alpha(\phi)$). Given input ϕ, let a machine M' dovetail[6] the algorithms to compute $M(\phi, i)$ for each $i \in \mathbb{N}$. Let M' output 0 if $1 \notin I_{M(\phi, i)}$ for some i,

[6] A machine can carry out a countable number of algorithms, in effect simultaneously, by "dovetailing" them. This means that the machine divides its time and tape space between the several algorithms. Construction of such machines is standard in the theory of computation.

and 1 if $0 \notin I_{M(\phi, i)}$ for some i. (These propositions can be decided effectively, given our choice of the sequence I.) Then M' $(\alpha, \nu_{\{0, 1\}})$-computes f (i.e., for all $\phi \in \mathrm{dom}(f \circ \alpha)$, $\nu_{\{0, 1\}} \circ M'(\phi) = f \circ \alpha(\phi)$).

$\Leftarrow$: First, construct a recursive sequence $\{\beta_i\}$ listing all rational intervals that contain 0 (i.e., $0 \in I_i$ iff for some j, $\beta_j = i$) and a similar sequence $\{\gamma_i\}$ for 1. Now suppose f is $(\alpha, \nu_{\{0, 1\}})$-computable as witnessed by M, i.e., for all $\phi \in \mathrm{dom}(f \circ \gamma)$, $\nu_{\{0, 1\}} \circ M(\phi) = f \circ \alpha(\phi)$. Let $M'(\phi, i) = \beta_i$ for all i if $M(\phi) = 0$ (in which case $\nu_{\{0, 1\}}(M(\phi)) = 0$), and let $M'(\phi, i) = \gamma_i$ if $M(\phi){\downarrow}$ and $M(\phi) > 0$ (in which case $\nu_{\{0, 1\}}(M(\phi)) = 1$). Then M' $(\alpha, \rho_{\mathbb{R}})$-computes f (i.e., for all $\phi \in \mathrm{dom}(f \circ \alpha)$, $\rho_{\mathbb{R}}(M'^{\phi}) = f \circ \alpha(\phi)$). ■

Corollary 3.2.14. $f\colon \subseteq \mathbb{R} \to \mathbb{R}$ is $(\rho_{\text{base-}b}, \nu_{\{0, 1\}})$-computable iff it is $(\rho_{\mathbb{R}}, \nu_{\{0, 1\}})$-computable.

Proof. Immediate from 3.2.11 and 3.2.13. ■

Hence, for concepts of decidability on $\mathbb{R}$ that are sufficiently related to characteristic functions, it makes no difference with which of the four representations of $\mathbb{R}$ defined above we choose to work.

3.2.4. The Topological Use Principle and the Neighborhood Enumeration Theorem

The Use Principle from classical recursion theory lies at the heart of several results in this dissertation. The Use Principle says that if a machine halts, it does so on the basis of some finite portion of the information provided to it.

Definition 3.2.15 (see Soare 1987). (i) Let $s = (s_0, s_1, \ldots, s_n) \in \mathbb{N}^n$. We say that length$(s) = n + 1$. For any $\phi \in \mathbb{N}^{\mathbb{N}}$, we write $s \subset \phi$ and $\phi \supset s$ if s is a *proper initial segment* of ϕ, i.e., if $s_i = \phi_i$ for all $i \in \{0, 1, \ldots, n\}$.

(ii) We write $M(s, i)\downarrow$ and $M(s, i) = j$ if for all $\phi \in \mathbb{N}^{\mathbb{N}}$, $\phi \supset s$ implies that $M(\phi, i) = j$. If this fails for every j, we write $M(s, i)\uparrow$.

Proposition 3.2.16 (Use Principle). $M(\phi, i) = j \Leftrightarrow (\exists s \subset \phi)\, M(s, i) = j$.

Proof. See Soare 1987. ■

In an uncountable topological space, finite information is approximate information. Hence, if a machine halts on x (i.e., given a standard representation of x), then it halts on an open neighborhood of x. (See Figure 2.2.) More precisely,

Proposition 3.2.17 (Topological Use Principle or TUP). If $M(\phi, a) = b \in \mathbb{N}$ for some $\phi \in \mathbb{N}^{\mathbb{N}}$ and $a \in \mathbb{N}$ such that $\rho_X(\phi) = x \in X$, then there is an open set $U \subseteq \tau_X$ containing x such that $(\forall y \in U)(\exists \psi \in \mathbb{N}^{\mathbb{N}})\, \rho_X(\psi) = y$ and $M(\psi, a) = b$.

Proof. Suppose x, ϕ, a, and b satisfy the antecedent. By the Use Principle, choose a finite sequence s such that $s \subset \phi$ and $M(s, a) = b$. Let $U = \bigcap_{i=0}^{\text{length}(s)-1} \sigma_{s_i}$. To see that U

satisfies the consequent, let $y \in U$. Choose θ such that $\rho_X(\theta) = y$ and let $\psi = s\theta$; that is,

$$\psi_i = \begin{cases} s_i & \text{if } i < \text{length}(s), \\ \theta_{i+\text{length}(s)} & \text{otherwise.} \end{cases}$$

Then $\rho_X(\psi) = y$. (The additive term 'length(s)' in the subscript on θ ensures that ψ includes all of θ and therefore lists *all* subbasic sets containing y, as required by the standard representation ρ_X.) By the Use Principle, $M(\psi, a) = M(s, a) = b$. ■

This principle is not purely an artifact of the standard representation, which represents points by sequences of open sets. It holds just as well, for example, under the strict regular Cauchy representation of real n-tuples (Parker 2003). It is equivalent to the fact that computable functions (with respect to reasonable, "admissible" representations) are continuous (Weihrauch 2000, 71).

The TUP is constructive and can be made "uniform," meaning that given any x where M halts and outputs b, we can effectively *construct* a neighborhood U of x where M outputs b. Furthermore, we can effectively enumerate the *entire* region where M outputs b, or that where M halts, as a recursive sequence of neighborhoods, each of which is a finite intersection of subbase sets. This will be useful in proving some positive decidability results.

To state the claim precisely, we first define enumerations of the finite strings and the finite intersections of subbase sets:

Definition 3.2.18. (i) For all $n, s_1, \ldots, s_n \in \mathbb{N}$, let $S(\langle n, s_1, \ldots, s_n \rangle) = (s_1, \ldots, s_n)$, and let $S(i)_j$ denote the j^{th} element of $S(i)$.

(ii) Relative to any coded space (X, σ), let $U_i = \bigcap_{j=0}^{\text{length}(S(i))-1} \sigma_{S(i)_j}$.

Intuitively, U_i is the set of points that have ρ_X-names beginning with the finite string $S(i)$.

Theorem 3.2.19 (Neighborhood Enumeration Theorem or NET). For any infinite-input Turing machine M, there is a recursive sequence $\{u_i\}$ such that $x \in \bigcup_i U_{u_i}$ iff for some $\phi \in \rho_X^{-1}(x)$, $M(\phi)\downarrow$.

Proof. We give an informal procedure to construct $\{u_i\}$. First, set $i = 0$. Then dovetail the computations of $M(S(j))$ over all $j \in \mathbb{N}$. Whenever M halts on some $S(j)$ without having scanned any input cells beyond length($S(j)$), set $u_i = j$ and increment i by one.

Now suppose $x \in \bigcup_i U_{u_i}$. Then for some i, x has a ρ_X-name ϕ beginning with $S(u_i)$. By construction of $\{u_i\}$ and the Use Principle, $M(\phi)\downarrow$. Conversely, suppose $\phi \in \rho_X^{-1}(x)$ and $M(\phi)\downarrow$. Again by the Use Principle and our construction of $\{u_i\}$, there is some i such that $S(u_i) \subset \phi$. Therefore, $x \in U_{u_i} \subseteq \bigcup_i U_{u_i}$. ■

3.3. Open and closed sets

In Weihrauch's Type-Two Theory of Effectivity, computability properties of sets are typically defined only for special subclasses of the Borel sets, such as the open sets.

We consider a few such notions in this section. Throughout, we assume X is the domain of a coded topological space and ρ_X the standard representation for X.

We begin with a notion that pre-dates TTE and is analogous to the classical notion of a recursively enumerable (r.e.) set.

Definition 3.3.1 (see Weihrauch 2000; cf. Lacombe 1957). A set $A \subseteq X$ is *r.e. open* if there is a machine M such that for all $\phi \in \mathrm{dom}(\rho_X)$, $M(\phi){\downarrow} \Leftrightarrow \rho_X(\phi) \in A$.

This is just the natural notion of a set such that some algorithm will tell us if an object is in the set, but says nothing otherwise. Such sets are also called *semirecursive* (Moschovakis 1980, Myrvold 1997) and Σ_1^0 (see Moschovakis 1980). The class Σ_1^0 is just one in the hierarchy of the Kleene pointclasses (ibid.). We will not be much concerned with the Kleene pointclasses, but those with superscript 0 are shown in our final diagram (Figure 3.3). They can be defined inductively as follows: For each $n \geq 1$, a Π_n^0 set is the complement of a Σ_n^0 set, a Δ_n^0 set is one that is both Σ_n^0 and Π_n^0, and a Σ_{n+1}^0 set is a union over some effective list of Π_n^0 sets.[7]

Notice that if there is a recursive enumeration of the subbase sets included in A, then A is r.e. open. That is,

[7] Hemmerling (2003) substitutes the superscript "ta" for zero and calls this the topological-arithmetic hierarchy.

Remark 3.3.2.[8] Assume (as always) that $\mathbf{X} = (X, \sigma)$ is a coded topological space. If $\{a_i\}$ is a recursive sequence of natural numbers and $A = \bigcup_{i=0}^{\infty} \sigma_{a_i}$ then A is r.e. open.

Proof. We give an algorithm for a machine M that halts on $\phi \in \mathrm{dom}(\rho_{\mathbf{X}})$ iff $\rho_{\mathbf{X}}(\phi) \in A$: Let M merely enumerate $\{a_i\}$ and, by dovetailing, check whether any ϕ_j equals some a_i. If so, let M halt. ■

In $\mathbb{R}^n$, the converse holds as well, and this justifies the remark made in Section 2.5 that r.e. open sets are enumerable in a very literal sense.

Proposition 3.3.3. $A \subseteq \mathbb{R}^n$ is r.e. open iff there exists a recursive sequence $\{a_i\} \in \mathbb{N}^{\mathbb{N}}$ such that $A = \bigcup_i I^n_{a_i}$.

Proof. $\Rightarrow$: Suppose A is r.e. open. Then by definition there is a machine M such that for all $x \in \mathbb{R}^n$ and all $\phi \in \rho_{\mathbb{R}^n}{}^{-1}(x)$, $M(\phi)\downarrow \Leftrightarrow x \in A$. By the Neighborhood Enumeration Theorem, there is a recursive sequence $\{u_i\}$ such that $x \in \bigcup_i U_{u_i}$ iff for some $\phi \in \rho_{\mathbf{X}}{}^{-1}(x)$, $M(\phi)\downarrow$, so $\bigcup_i U_{u_i} = A$. In this case, each $U_i = \bigcap_{j=0}^{\mathrm{length}(S(i))-1} I^n_{S(i)_j}$. The set of open rational n-intervals I^n_i is closed under finite intersection, so each $U_i \in I^n$. Define $f\colon \mathbb{N} \to \mathbb{N}$ by $f(i) = j \Leftrightarrow U_i = I^n_j$. By construction of $\{U_i\}$ and I^n, f is recursive. Now just let $\{a_i\} = \{f(u_i)\}$. Then $\{a_i\}$ is recursive and $\bigcup_i I^n_{a_i} = \bigcup_i U_{u_i} = A$.

[8] This is a well known triviality.

$\Leftarrow$: By Remark 3.3.2. ■

This also shows that in $\mathbb{R}^n$, r.e. open sets are really open. Comfortingly, the same is true for all coded topological spaces.

Remark 3.3.4.[9] All r.e. open (i.e., semirecursive) sets are open.

Proof. Suppose A is r.e. open. Then there is a machine M such that for all $\phi \in \mathrm{dom}(\rho_X)$, $M(\phi)\downarrow \Leftrightarrow \rho_X(\phi) \in A$. By the Topological Use Principle, for each $x \in A$ there is an open set $U(x)$ such that $x \in U(x)$, and for each $y \in U(x)$ a sequence $\psi \in \mathbb{N}^{\mathbb{N}}$ such that $\rho_X(\psi) = y$ and $M(\psi)\downarrow$. Hence each such y is in A, so each $U(x) \subseteq A$. Therefore $A = \bigcup_{x \in A} U(x)$, so A is open. ■

Equivalently, the domain of any computable function $f: X \to \mathbb{N}$ is an open set. This can also be seen as a consequence of another fact (which we will not prove): that the domain of a computable function taking infinite strings to finite strings is always open in the natural topology for infinite strings (Weihrauch 2000, 31). Since here we are concerned with computation conceived as the manipulation of strings, facts about computation on strings (or equivalently, on natural numbers and sequences thereof) underlie all of our results.

Given the notion of an r.e. open set, a natural definition of a decidable set might

[9] This is also well known; Myrvold mentions it for example in 1997.

be as follows:

Definition 3.3.5 (see Myrvold 1997). A set $A \subseteq X$ is *naïvely decidable* if A and A^C are r.e. open.

Equivalently, a set is naively decidable if there is a machine that computes its characteristic function, for if there is a machine that halts on all and only elements of A and another that does the same for A^C, we can construct from these a third machine that gives output 1 if the input is in A and 0 otherwise. In any case, naïve decidability is rather special, for,

Remark 3.3.6. If $A \subseteq X$ is naïvely decidable then A is open and closed.

Proof. Immediate from 3.3.5. ■

In a connected space **X**, the only sets that are both open and closed are X and $\varnothing$. So for example, the only naïvely decidable sets in $\mathbb{R}^n$ are $\mathbb{R}^n$ and $\varnothing$.

There is also a notion of an r.e. closed set in the literature. Weihrauch (2000) defines an r.e. closed set in $\mathbb{R}^n$ as a closed set such that we can effectively list all the open rational cubes that intersect it. More generally,

Definition 3.3.7 (Weihrauch; cf. Zhou 1996). A closed set $A \subseteq X$ is *r.e. closed* if there exists a recursive sequence $\{a_i\}_{i \in \mathbb{N}}$ of natural numbers such that $j = a_i$ for some i if and only if $\sigma_j \cap A \neq \varnothing$.

Since the complement of an open set is closed and vice-versa, it seems fairly natural to define recursive open and closed sets as follows:

Definition 3.3.8 (Weihrauch). (i) An open set is *recursive* if it is r.e. open and its complement is r.e. closed.

(ii) A closed set is *recursive* if it is r.e. closed and its complement is r.e. open.

(iii) A recursive open or closed set is called a *TTE-recursive* set.

Notice that a set is TTE-recursive if and only if its complement is TTE-recursive.

This definition is partly justified by the idea that a distance function $d_A(x) = \inf_{a \in A} d(x, a)$, where d is some metric on X, is a natural generalization of the notion of a characteristic function. It turns out that,

Proposition 3.3.9 (Weihrauch). A closed set $A \subseteq \mathbb{R}^n$ is TTE-recursive iff $d_A(x)$ is computable, where $d_A(x) = \inf_{a \in A} \|x - a\|$ and $\|\cdot\|$ is the usual, Euclidean norm.

Proof. See Weihrauch 2000 (128). ■

However, the notions of a recursive closed set and a computable distance function also have certain counterintuitive asymmetries, as discussed in Chapter 2. We can effectively determine when a given point is *not* in a recursive closed set A, but no procedure will determine when a point *is* in A, nor even, in general, when a point is in the *interior* of A. Zhou (1996) offers a stronger, more symmetric notion:

Definition 3.3.10 (Zhou 1996). An open or closed set A is *strongly recursive* if $\mathrm{int}(A)$ and $\mathrm{ext}(A)$ are recursive open sets.

Hertling calls this *birecursiveness* (1999). Note that if A is strongly recursive, then *a fortiori*, $\mathrm{int}(A)$ and $\mathrm{ext}(A)$ are r.e. open. Also, a set is strongly recursive if and only if its complement is strongly recursive. Incidentally, strong recursiveness also implies that if $\mathbb{N} \cap \partial A = \varnothing$ then $\mathbb{N} \cap A$ is recursive. This is not the case for recursive open or closed sets in general (see for example Hemmerling 2003, Example 1), so strong recursiveness is indeed stronger than TTE-recursiveness.

3.4. Topological notions of decidability for arbitrary sets

A more symmetric generalization of the notion of a characteristic function is that of a *Δ-function:*

Definition 3.4.1 (Hemmerling 2003). (i) The *Δ-function* for a set $A \subseteq \mathbb{R}^n$ is $\Delta_A(x) = d_A(x) - d_{X\setminus A}(x)$. (See Proposition 3.3.9.)

(ii) A set $A \subseteq \mathbb{R}^n$ is *Δ-decidable* if its Δ-function is computable.

Hertling (1999) considers this notion for closed sets while Hemmerling (2003) extends it to arbitrary subsets of $\mathbb{R}^n$. Hertling shows that

Proposition 3.4.2. If $A \subseteq \mathbb{R}^n$ is closed then A is Δ-decidable iff A is strongly recursive.

Proof. See Hertling 1999. ■

Corollary 3.4.3.[10] If $A \subseteq \mathbb{R}^n$ is open then A is Δ-decidable iff A is strongly recursive.

Proof. Suppose A is open. Then A^C is closed, so A is Δ-decidable $\Leftrightarrow A^C$ is Δ-decidable $\Leftrightarrow A^C$ is strongly recursive $\Leftrightarrow A$ is strongly recursive. ■

Therefore,

Proposition 3.4.4 (Hemmerling 2003). If $A \subseteq \mathbb{R}^n$ is strongly recursive then A is Δ-decidable.

Proof. Immediate from 3.4.2 and 3.4.3. ■

The converse clearly fails simply because strongly recursive sets must be open or closed.

[10] Another well known triviality.

For example, the interval [0, 1) is Δ-decidable but not strongly recursive; for that matter, it is not even r.e. open or r.e. closed.

Myrvold (1997) introduced another natural notion of decidability for arbitrary sets:

Definition 3.4.5 (Myrvold 1997; cf. Penrose 1989). A set $A \subseteq X$ is *decidable ignoring boundaries* (d.i.b.) if int(A) and ext(A) are r.e. open.

This implies the existence of a decision procedure that succeeds everywhere except on the boundary, and furthermore *does not halt* on the boundary. Thus the procedure must not so much *ignore* the boundary as tiptoe carefully around it. Penrose dismisses an intuitive notion similar to d.i.b. because the complexity of a set like the Mandelbrot set seems to reside in the shape of its border. Hence a computability concept that ignores borders will not reflect such complexity. However, d.i.b. does not ignore borders in that sense; if the boundary of a set is too complex for an algorithm to tiptoe around, the set is not d.i.b.

Notice that Δ-decidability implies d.i.b. Hemmerling shows that,

Proposition 3.4.6. For any set $A \subseteq \mathbb{R}^n$, A is Δ-decidable iff int(A) and ext(A) are r.e. open and ∂A is r.e. closed.

Proof. See Hemmerling 2003. ■

Hence it follows immediately that,

Corollary 3.4.7. If $A \subseteq \mathbb{R}^n$ is Δ-decidable then A is d.i.b.

There is no reason to expect the boundary of a d.i.b. set to be r.e. closed, so it is not surprising that the converse of Corollary 3.4.7 fails, as we see next.

Proposition 3.4.8. There exists a set $A \subseteq \mathbb{R}$ that is d.i.b. but not Δ-decidable.

Proof. Let K be any non-recursive r.e. subset of $\mathbb{N}$. (A classic example is the set $K = \{e : \varphi_e(e)\downarrow\}$, where $\{\varphi_e\}$ is some recursive enumeration of the partial recursive functions on $\mathbb{N}$.) Let $\{k_i\}_{i \in \mathbb{N}}$ be a recursive enumeration of K without repetition and let $a_i = \sum_{j=0}^{i} 2^{-k_j}$ for each i. Finally, let $A = \mathbb{Q} \cap (\lim \{a_i\}, \infty)$. Then $\operatorname{int}(A) = \varnothing$, and $\operatorname{ext}(A) = (-\infty, \lim \{a_i\})$. Both are r.e. open: the interior trivially, and the exterior because we can construct a machine M that, given input ϕ, enumerates the sequence $\{a_i\}$ and halts if for some i and j, $I_{\phi_j} \subseteq (-\infty, a_i)$. Therefore A is d.i.b.

To see that A is not Δ-decidable, suppose it is. We will construct an algorithm to tell us whether or not a given n is in K, contradicting the non-recursiveness of K. First, notice that, since A is Δ-decidable, Δ_A is computable, so d_A is computable. But $d_A = d_{\partial A}$, so this implies that $\partial A = [\lim \{a_i\}, \infty)$ is recursive closed and therefore r.e. closed. That means we have a recursive enumeration $\{b_i\}$ of all open rational intervals intersecting ∂A. Note also that $n \notin K$ iff for some j, two things hold: n is not any of the first j entries in the enumeration $\{k_i\}$, and $a_j + 2^{-n} > \lim \{a_i\}$. So to decide whether $n \in K$, we need only enumerate $\{k_i\}$, $\{a_i\}$, and $\{b_i\}$, checking for each j whether $n = k_j$ (in which case $n \in K$) and whether, instead, $n \neq k_i$ for any $i < j$ but $b_j = (0, a_j + 2^{-n})$ (in which case $n \notin K$). Therefore K is recursive, contrary to assumption. ■

Finally, note that neither recursive open nor recursive closed implies d.i.b., since, as Zhou shows (1996; see also Hemmerling 2003), there are recursive closed sets that are not strongly recursive, i.e., their interiors are not r.e. open. Hence they are not d.i.b. Their complements are of course recursive open, and still not d.i.b.

3.5. Measure-theoretic decidabilities

Ko (1991) also considers a notion of decidability for arbitrary sets. He calls a set *recursively approximable* if it can be correctly decided up to arbitrarily small Lebesgue measure. We state this precisely while generalizing it to arbitrary coded metric spaces and measures (or outer measures).

Definition 3.5.1 (Ko). Let μ be a measure (or outer measure) on X. A set $A \subseteq X$ is *μ-recursively approximable* (μ-r.a.) if there is a machine M such that for all $x \in X$, $n \in \mathbb{N}$, and $\phi \in \mathrm{dom}(\rho_X)$,

(i) if $\rho_X(\phi) = x$ then $M(\phi, n)\downarrow$, and

(ii) if $n > 0$ then $\mu E_{A,\,n}(M) \leq 2^{-n}$, where $E_{A,n}(M) = \{x \in X \mid [\exists \phi \in \mathrm{dom}(\rho_X)]\ M(\phi, n) \neq \chi_A(x)\}$, called the n^{th} *error set for M.*

Intuitively, $E_{A,n}(M)$ is the set of points where M miscalculates the membership of A given input n.

Ko [2] shows that on the interval [0, 1] with Lebesgue measure λ, λ-r.a. is equivalent to Šanin's notion of recursive measurability (Šanin 1968). This will be useful below, so we state the definition and result. We use $\mathbb{D}$ to denote the set of *dyadic rational numbers* $a/2^n$, where $a \in \mathbb{Z}$ and $n \in \mathbb{N}$. The set of finite sequences of dyadic rationals is written $\mathbb{D}^*$.

Definition 3.5.2 (Šanin 1968). (i) A sequence $\{S_i\}$ of sets in $\mathbb{R}$ is a *recursive sequence of sets* if there is a recursive function $\varphi: \mathbb{N} \to \mathbb{N} \times \mathbb{D}^*$ such that if $\varphi(i) = (k, (a_1, b_1, \ldots, a_k, b_k))$ then $a_1 < b_1 \leq a_2 < b_2 \leq \ldots \leq a_k < b_k$ and $S_i = \bigcup_{j=1}^{k}(a_j, b_j)$.

(ii) A set $S \subseteq \mathbb{R}$ is *recursively measurable* (*r.m.*) if there exists a recursive sequence $\{S_i\}$ of sets in $\mathbb{R}$ such that for all $i > 0$, $\lambda(S \,\Delta\, S_i) \leq 2^{-i}$, where Δ is the symmetric difference operator.

Theorem 3.5.3 (Ko). A set $S \subseteq [0, 1]$ is λ-r.a. if and only if it is r.m.

Proof. See Ko 1991 (162). ■

One might expect that recursively open and closed sets should be recursively measurable, in parallel with classical analysis. However, the notions are independent. In Section 6 we will meet a subset of [0, 1] that is closed and strongly recursive, and hence TTE-recursive, but not λ-r.a. and hence not recursively measurable. The fact that λ-r.a. (and hence r.m.) does not imply TTE-recursiveness will follow from 3.5.7, 3.5.8, and 3.5.10.

One can demand more of a set than to be recursively approximable. One might well prefer a decision procedure that works not only on all but an arbitrarily small set, but on all but a measure-zero set. Also, as long as we permit errors on a measure-zero set, we might also permit the decision procedure not to halt on some measure-zero set. This leads us to the following definition of *decidability up to μ-measure zero.*

Definition 3.5.4. For a given measure (or outer measure) μ on X, a set $A \subseteq X$ is *decidable up to μ-measure zero* (or *decidable mod zero,* abbreviated d.m.z. or μ-d.m.z.) if there exists a machine M such that $\mu E_A(M) = 0$, where $E_A(M)$ is the *error set for M,* defined by $E_A(M) = \{x \in X \mid [\exists \phi \in \rho_X^{-1}(x)]\ [M(\phi)\uparrow \text{ or } M(\phi) \neq \chi_A(x)]\ \}$.

Neither μ-d.m.z. nor μ-r.a. is completely trivial in extension. For example,

Remark 3.5.5. If $\mu A = 0$ or $\mu A^C = 0$ then A is μ-d.m.z. and μ-r.a.

Proof. If $\mu A = 0$, let $M^{\phi}(n) \equiv 0$ for all ϕ, n. Then $E_A(M) = E_{A,n}(M) = A$, so $\mu E_A(M) = \mu E_{A,n}(M) = 0$. If instead $\mu A^C = 0$, let $M^{\phi}(n) \equiv 1$. ■

Later we will see other examples of μ-r.a. and μ-d.m.z. sets that have non-zero measure and complements of non-zero measure.

Even in $\mathbb{R}$ with Lebesgue measure λ, d.m.z. does not imply d.i.b.:

Proposition 3.5.6. There exists a set in $\mathbb{R}$ that is λ-d.m.z. (and λ-r.a.) but not d.i.b.

Proof. By 3.5.5, all we need is a measure-zero set that is not d.i.b. Let K be any non-recursive r.e. subset of $\mathbb{N}$. Embedded in $\mathbb{R}$, K has Lebesgue measure zero, so K is trivially λ-r.a. and λ-d.m.z. Now suppose toward contradiction that K is d.i.b. This implies that $\mathrm{ext}(K) = \mathbb{R} \setminus K$ is r.e. open. Since $\rho_{\mathbb{R}} \equiv \rho_C$, this implies that there is a machine M such that if $\phi \in \rho_C^{-1}(x)$ then $M(\phi)\downarrow \Leftrightarrow x \in \mathbb{R} \setminus K$. For each $n \in \mathbb{N}$, define $\phi^n \in \rho_C^{-1}(n)$ by $\phi^n(i) \equiv n$. Now it is trivial to construct a machine M' such that $M'(n)\downarrow$ if and only if $M(\phi^n)\downarrow$. Such an M' halts on an input $n \in \mathbb{N}$ if and only if $n \notin K$. Hence $\mathbb{N} \setminus K$ is r.e., but since K is r.e., this contradicts the non-recursiveness of K. ■

This implies that λ-d.m.z. does not imply Δ-decidability nor strong recursiveness. (From here on, it may be helpful to refer occasionally to Figure 3.3.)

We will see later that the converse fails as well; strong recursiveness, and therefore d.i.b. and Δ-decidability, do not imply d.m.z., even in the context of $\mathbb{R}$ with

Lebesgue measure (3.6.3). However, for sets with trivial boundaries, d.i.b. does imply d.m.z.:

Remark 3.5.7. If $\mu\partial A = 0$ and $A \subseteq X$ is d.i.b. then A is μ-d.m.z.

Proof. If int(A) and ext(A) are r.e. open, then one easily constructs a machine M that gives output 1 on input $\phi \in \rho_X^{-1}(x)$ iff $x \in$ int(A) and 0 iff $x \in$ ext(A). Hence $E_A(M) = \partial A$, so $\mu E_A(M) = 0$. ■

For the sake of completeness, we now show that the property being d.i.b. with a measure-zero boundary does not imply Δ-decidability, nor TTE-recursiveness, nor, therefore, strong recursiveness.

Proposition 3.5.8. There exist both open d.i.b. sets $A \subseteq \mathbb{R}$ and closed d.i.b. sets $A \subseteq \mathbb{R}$ such that $\lambda\partial A = 0$ and A is not Δ-decidable, nor TTE-recursive.

Proof. Choose a non-recursive r.e. set $K \subseteq \mathbb{N}$. Let $A = (\mathbb{N} \setminus K) \subseteq \mathbb{R}$. Then A^C is r.e. open by Remark 3.3.2, for to enumerate the rational open intervals contained in A^C, it is enough to enumerate those contained in intervals $(i, i + 1)$ for $i \in \mathbb{Z}$, as well as those rational intervals J such that $\mathbb{N} \cap J \subseteq K$. Since int($A$) $= \varnothing$, A is d.i.b. However, d_A is not computable, for suppose it is. Then we can use it to effectively decide whether or not a given whole number is in K: if $n \in \mathbb{N}$ and $d_A(n) > 0$ then $n \in K$, while if $n \in \mathbb{N}$ and $d_A(n)$

< 1 then $n \notin K$. This contradicts the non-recursiveness of K, so d_A is not computable. Therefore A is neither TTE-recursive nor Δ-decidable, and the same goes for A^C. ■

This also shows that d.m.z. does not imply Δ-decidability, nor does it imply TTE-recursiveness. (The latter point is trivial, since TTE-recursiveness holds only for open and closed sets.)

We now compare d.m.z and r.a. For an arbitrary measure μ, μ-d.m.z. and μ-r.a. are independent. We show first by example that even a strongly recursive interval may be μ-d.m.z. but not μ-r.a.

Proposition 3.5.9. There exists a measure μ on $\mathbb{R}$ such that the interval $B = [0, \infty)$ is μ-measurable and μ-d.m.z. but not μ-r.a.

Proof. For any Lebesgue-measurable set $A \subseteq \mathbb{R}$, let $\mu A = \int_A |1/x|\, dx$. Since $\mu\partial B = \mu\{0\} = 0$, B is μ-d.m.z. To see that B is not μ-r.a., suppose some machine M halts on every input $(\phi, n) \in \rho_{\mathbb{R}}^{-1}(x) \times \mathbb{N}$, $x \in \mathbb{R}$. Fix n. By the Topological Use Principle (3.2.17), fix $\psi \in \rho_{\mathbb{R}}^{-1}(0)$ and a neighborhood U of zero such that $(\forall x \in U)[\exists\phi \in \rho_{\mathbb{R}}^{-1}(x)]\ M(\phi, n) = M(\psi, n)$. (See Figure 3.2.) Then there exists and interval $(-a, a) \subseteq U$. But either $(-a, 0)$ or $(0, a) \subseteq E_{B,n}(M)$. By choice of μ, $\mu(-a, 0) = \mu(0, a) = \infty$. Therefore $\mu E_{B,n}(M) = \infty$. Hence there is no machine M witnessing that B is μ-r.a. ■

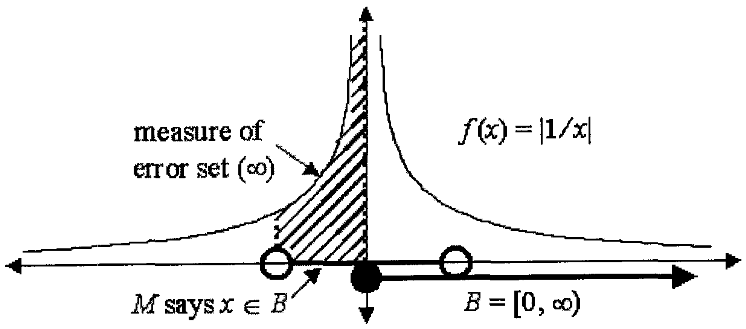

Figure 3.2. Proof of Proposition 3.5.9.

Note that the d.m.z. set in this proof, $[0, \infty)$, is of course strongly recursive: $(-\infty, 0)$ and $(0, \infty)$ are r.e. open, and $(-\infty, 0]$ and $[0, \infty)$ are r.e. closed.

Here we have used a strange measure μ to show that μ-d.m.z. (and strong recursiveness) does not imply μ-r.a. in the general case. In $\mathbb{R}^n$ with Lebesgue measure λ, this is not the case; λ-d.m.z. does imply λ-r.a. (3.5.10), but we will see in the next section that λ-r.a. does not imply λ-d.m.z.

We now show that λ-d.m.z. implies λ-r.a. The difficulty in this lies with the requirement in the definition of that the approximation procedure must halt everywhere. Our strategy is this: Given an algorithm M that shows a set $A \subseteq \mathbb{R}^n$ is d.m.z., and given a desired error bound 2^{-m}, we construct two recursive sequences of neighborhoods: the neighborhoods U_{u_i} (of the Neighborhood Enumeration Theorem 3.2.19), where M halts and which nearly fill $\mathbb{R}^n$, and some very small neighborhoods V_{v_i} covering the gaps, where our output will be arbitrary. We construct $\{V_{v_i}\}$ in stages, first covering the gaps within the n-cube $[-1, 1]^n$, then those in $[-2, 2]^n$, etc., with the new additions to $\{V_{v_i}\}$ shrinking rapidly with each stage. Thus the total measure of $\{V_{v_i}\}$ can be made arbitrarily small. An illustrative example follows the proof.

Theorem 3.5.10. If $A \subseteq \mathbb{R}^n$ is λ-d.m.z. then A is λ-r.a.

Proof. Recall $\{S(i)\}$ and $\{U_i\} = \{U(i)\}$ of Definition 3.2.18, and for each $i \in \mathbb{N}$, let $V(i) = \bigcup_{j=0}^{\text{length}(S(i))-1} I^n_{S(i)_j}$.

Now suppose A is d.m.z. as witnessed by a machine M. By the Neighborhood Enumeration Theorem (3.2.19), there is a recursive sequence $\{u_i\}$ such that $x \in \bigcup_i U(u_i)$ iff for some $\phi \in \rho_X^{-1}(x)$, $M(\phi)\downarrow$. For each $m \in \mathbb{N}$, construct a recursive sequence of integers $\{v_{m,i}\}_{i \in \mathbb{Z}^+}$ as follows:

1. Set $i = 1$.
2. Find and fix j (temporarily) such that $\lambda([-i, i]^n \setminus \bigcup_{k=0}^{j} U(u_k)) < 2^{-m-i}$.
3. Find and fix $v_{m,i}$ such that $\lambda V(v_{m,i}) < 2^{-m-i}$ and $V(v_{m,i}) \cup \bigcup_{k=0}^{j} U(u_k)$ covers $[-i, i]^n$. Set $i = i + 1$ and repeat from step 2.

Step 2 is possible because $\lambda[\mathbb{R}^n \setminus \bigcup_{k=0}^{\infty} U(u_k)] = \lambda\{x \in \mathbb{R}^n\text{: } [\forall \phi \in \rho_X^{-1}(x)]\ M(\phi)\uparrow\} = 0$, and volumes of finite unions, intersections, and differences of rational n-intervals can be computed exactly. Step 3 is possible because the intervals $[-i, i]$ are compact. Note also $\bigcup_i [U(u_i) \cup V(v_{m,i})] = \mathbb{R}^n$, and $\lambda \bigcup_i V(v_{m,i}) < \sum_i 2^{-m-i} = 2^{-m}$.

Now an algorithm for a machine M' satisfying Definition 3.5.1 (r.a.) is simple: given input $(\phi, m) \in \rho_{\mathbb{R}^n}^{-1}(x) \times \mathbb{N}$, evaluate for each i whether $x \in U(u_i)$ and whether $x \in$

$V(v_{m,i})$. If $x \in U(u_i)$, output $M(S(u_i))$. If $x \in V(v_{m,i})$, output zero.

Since $\bigcup_i [U(u_i) \cup V(v_{m,i})] = \mathbb{R}^n$, M' will halt on every $\rho_{\mathbb{R}^n}$-name for a point in $\mathbb{R}^n$. Also, $E_{M',m}(A) \subseteq E_M(A) \cup \bigcup_i V(v_{m,i})$. Since $\lambda E_M(A) = 0$ and $\lambda \bigcup_i V(v_{m,i}) < 2^{-m}$, $\lambda E_{M',m}(A) < 2^{-m}$. ■

Some informal discussion of an example may help to clarify the above proof.

Example 3.5.11. Let A be the set of real numbers such that the closest integer is unique and prime, i.e., $A = (3/2, 5/2) \cup (5/2, 7/2) \cup (9/2, 11/2) \cup \ldots$. This set is d.m.z. and also d.i.b., so suppose M decides it correctly except on the boundary $\{3/2, 5/2, 7/2, \ldots, p_i - 1/2, p_i + 1/2, \ldots\}$, and that M does not halt on that boundary. We wish to see how the machine M' of the preceding proof might proceed, given input $(\phi, m) \in \rho_{\mathbb{R}}^{-1}(x) \times \mathbb{N}$,

First, M' begins constructing the sequence $U(u_i)$ of open rational intervals where M halts (that is, it computes the *indices* of these intervals). When these fill or nearly fill the interval $[-1, 1]$, M' constructs a finite string $v_{m,1}$ representing a very small open set $V(v_{m,1})$ that covers any gaps within $[-1, 1]$ that the sets $U(u_i)$ have not yet filled. Here $V(v_{m,1})$ will likely be empty, since M halts everywhere on $[-1, 1]$. Note that this construction is independent of the input ϕ.

At this point M' might check whether $\rho_{\mathbb{R}}(x)$ lies in either $U(u_1)$ or $V(v_{m,1})$. This is done in finite time by checking whether any interval $\bigcap_{i=0}^{j} I_{\phi_i}$ is a subset of $U(u_1)$ or $V(v_{m,1})$. In the former case, M' then applies the algorithm of M to $S(u_1)$ and outputs the result. In the latter case, $\rho_{\mathbb{R}}(\phi)$ lies within a permissibly small error set, so M' just outputs zero, willy-nilly.

If neither case holds, then M' proceeds on, counting the sets $U(u_i)$ needed to fill or nearly fill [-2, 2], and constructing an even smaller set $V(v_{m,2})$ covering $[-2, 2] \setminus \bigcup_i U(u_i)$. In the present example, M will not halt on 3/2, so the set $V(v_{m,2})$ will likely consist of one small interval such as $(3/2 - 2^{-m-4}, 3/2 + 2^{-m-4})$. M' then checks to see whether some $\bigcap_{i=0}^{j} I_{\phi_i}$ lies entirely within $U(u_2)$ or $V(v_{m,2})$, and if so, generates output accordingly. If not, these steps repeat until, for some j, $\bigcap_{i=1}^{j} I_{\phi_i}$ is found to lie entirely within some $U(u_i)$ or $V(v_{m,i})$.

To conclude this section, we introduce localized and globalized versions of μ-d.m.z. These will provide a sense in which the behavior of a *particular* system in a particular state might be called undecidable, to be discussed in subsequent chapters.

Definition 3.5.12. (i) A set $A \subseteq X$ is *decidable up to μ-measure zero* (μ-d.m.z.) *at a point* $x \in X$ if there exists a machine M and an open set $U \in \tau_X$ such that $x \in U$ and $\mu[E_A(M) \cap U] = 0$.

(ii) A is *everywhere μ-d.m.z.* if A is μ-d.m.z. at every point $x \in X$.

(iii) A is *somewhere μ-d.m.z.* if A is *μ-d.m.z.* at some point $x \in X$. Otherwise, A is *nowhere μ-d.m.z.*

Clearly μ-d.m.z. implies everywhere μ-d.m.z. and somewhere μ-d.m.z. However, the converses fail. Any set with an open subset is somewhere μ-d.m.z. (if μ is regular) but not necessarily μ-d.m.z.; we will see examples in the next section. Also, everywhere d.m.z. as defined above does not imply d.m.z. *simpliciter*, for,

Proposition 3.5.13. There exists a set $A \subseteq \mathbb{R}$ that is everywhere λ-d.m.z. but not λ-d.m.z.

Proof. Let K be any non-recursive set of integers and let $A = \bigcup_K [k, k+1]$. Then for any $x \in \mathbb{R}$ and any bounded open interval U containing x, $U \cap A$ consists of a finite union of intervals with integer endpoints, so clearly there exists a machine M_U such that $\lambda[E_A(M_U) \cap U] = 0$. Now suppose toward contradiction that A is λ-d.m.z. Then there is some machine M such that $\lambda E_A(M) = 0$, so for each $i \in \mathbb{Z}$ there is some $x \in (i, i+1)$ and $\phi \in \rho_{\mathbb{R}}^{-1}(x)$ such that $M(\phi) = \chi_A(x)$. Also, for any non-integer x and $\phi \in \rho_{\mathbb{R}}^{-1}(x)$, if $M(\phi)\downarrow$ then $M(\phi) = \chi_A(x)$ (for suppose not; then by the Topological Use Principle, there is some open set V and a $\psi \in \rho_{\mathbb{R}}^{-1}(y)$ for each $y \in V$ such that $M(\psi)\downarrow$ but $M(\psi) \neq \chi_A(y)$, and therefore $\lambda E_A(M) > 0$, contrary to hypothesis). Therefore, in order to decide whether $n \in \mathbb{N}$, one need only look for a finite initial segment r of a name $\phi \in \rho_{\mathbb{R}}^{-1}(x)$ for some $x \in (n, n+1)$ such that $M(r)\downarrow$. One will eventually turn up, and when it does, $\chi_K(n) = M(r)$. But this contradicts the non-recursiveness of K. ∎

3.6. Riddled sets

The following definition derives from work in non-linear dynamics (Alexander et al. 1992, Sommerer and Ott 1993, 1996). Intuitively, a set A is "riddled" with holes if it has holes—that is, portions of A^C—with positive measure in every neighborhood of A. More precisely,

Definition 3.6.1 (Alexander et al. 1992). A set $A \subseteq X$ is *riddled* (with respect to a measure or outer measure μ) if for every open set $U \subseteq X$, $\mu(U \setminus A) > 0$.

Note that for any *regular* measure μ, i.e., one that assigns positive value to all open sets, nowhere dense sets are always riddled with respect to μ. The converse fails; a riddled set might be dense, for its positive-measure "holes" need not be open sets. However, a riddled set is never r.e. open because it is not open.

Because of the topological use principle, riddled sets are not d.m.z. unless they have measure zero:

Theorem 3.6.2. If $A \subseteq X$ is riddled and $\mu A > 0$, then A is not μ-d.m.z.

Proof. Assume the antecedent, and suppose toward contradiction that A is μ-d.m.z. and $\mu E_A(M) = 0$. Since $\mu A > 0$, there is at least one $x \in A$ such that M halts and outputs 1 given some input $\phi \in \rho_X^{-1}(x)$. Thus the set $B = \{x \mid [\exists \phi \in \rho_X^{-1}(x)]\ M(\phi) = 1\}$ is non-empty. By the Topological Use Principle, B is open. Because A is riddled, $\mu(B \setminus A) > 0$.

But for every $x \in B$, there is some $\phi \in \rho_X^{-1}(x)$ such that $M(\phi) = 1$, so $B \setminus A \subseteq E_A(M)$. Therefore $\mu E_A(M) > 0$, contrary to hypothesis. ■

This result makes it possible to show that certain interesting sets of states in dynamical systems are not d.m.z. (See Chapters 4 and 5.) It also enables us to add the following result to our study of the logical relations among decidability concepts:

Proposition 3.6.3. There exists a set $C \subseteq \mathbb{R}$ that is λ-r.a. (and, by the way, strongly recursive) but not λ-d.m.z.

Proof. Say a closed interval J is *maximal* in a set $S \subseteq \mathbb{R}$ if $J \subseteq S$ and for every closed interval $K \subseteq S$, $K \cap J \neq \varnothing \Rightarrow K \subseteq J$. (For closed S, the maximal intervals are just the connected components.) For all $i \in \mathbb{Z}^+$ let $r_i = 2^{-i-1}$. We construct a generalized Cantor set C as follows:

(i) Let $C_0 = [0, 1]$.

(ii) For each $i \in \mathbb{Z}^+$, let C_i =

$$C_{i-1} \setminus \bigcup_{a,b:[a,b] \text{ is maximal in } C_{i-1}} \left(\tfrac{a+b}{2} - 2^{-i-1} r_i, \tfrac{a+b}{2} + 2^{-i-1} r_i\right).$$

(iii) Let $C = \bigcap_i C_i$.

Part (ii) dictates that we obtain C_i by removing the middle segment of length $r_i / 2^i$ from each of the 2^i maximal intervals in C_{i-1}. Hence the limit of this process, the set C, has measure $1 - \Sigma r_i = 1/2$. Note C is riddled by construction, so C is not λ-d.m.z.

We show now that C is λ-r.a. Given input $(\phi, m) \in \mathbb{N}^{\mathbb{N}} \times \mathbb{N}$, let M examine the intervals I_{ϕ_i} in turn. Let M output 1 if it finds some maximal interval $[a, b]$ in C_{m+1} and some $I_{\phi_i} \subseteq (a - 2^{-2m+3}, b + 2^{-(2m+3)})$, but let M instead output 0 if it first finds some maximal interval $[a, b]$ in $\mathrm{cl}(C^{\mathrm{C}}_{m+1})$ and some $I_{\phi_i} \subseteq (a - 2^{-2m+3}, b + 2^{-(2m+3)})$. (This can be done effectively since one can construct the rational endpoints of C_{m+1} following (i)-(iii), and we assume that the enumerations $\{I_i\}$ of the rational intervals and $\{q_i\}$ of the rationals are recursive.) Then any $x \in E_{M,m}(C_{m+1})$ must be close to one of the 2^{m+2} endpoints of C_{m+1}, within distance $2^{-(2m+3)}$. Therefore $\lambda E_{M,m}(C_{m+1}) \leq (2^{m+2})(2^{-(2m+3)}) = 2^{-(m+1)}$. So, if $\lambda(C_{m+1} \setminus C) \leq 2^{-(m+1)}$, then $\lambda E_{M,m}(C) \leq 2^{-(m+1)} + 2^{-(m+1)} = 2^{-m}$.

This is in fact the case, for we now show that for all i, $\lambda(C_i \setminus C) < 2^{-i}$. For $i = 0$ we have $\lambda(C_0 \setminus C) = \lambda C_0 - \lambda C < 1 = 2^0$, since $\lambda C > 0$. For $i > 0$ we have $\lambda(C_i \setminus C) = \sum_{j=i+1}^{\infty} \lambda(C_{j-1} - C_j) =$ (by construction) $\sum_{j=i+1}^{\infty} 2^{-j} \lambda C_{j-1} < \sum_{j=i+1}^{\infty} 2^{-j} = 2^{-i}$.

Therefore $\lambda E_{M,m}(C) \leq 2^{-m}$, so C is r.a.

The proof that C is strongly recursive is exactly as in the proof of Proposition 3.6.6 below (where it is more important). ■

Incidentally, it follows that C^{C} of the proof is not d.m.z., though since C^{C} contains an open set, it is somewhere d.m.z. This confirms the trivial fact that a somewhere d.m.z. set is not necessarily d.m.z.

To complete our study of logical relations among decidabilities, we exhibit a set that is strongly recursive, and hence TTE-recursive, Δ-decidable, and d.i.b, but not λ-r.a. This is done by varying the construction of the above set C so that λC is a non-computable number. We will use two lemmas:

Lemma 3.6.4 (see Ko 1991). If $S \subseteq \mathbb{R}$ is r.m. then λS is a computable number.

Proof. Assuming S is r.m., let $\{S_n\}$ be a recursive sequence of sets such that for all $n \in \mathbb{N}$, $\lambda(S \,\Delta\, S_n) \leq 2^{-n}$. By Definition 3.5.2 there is a recursive sequence that gives the endpoints $a_1, b_1, \ldots, a_{k(n)}, b_{k(n)}$ of each S_n. Let $M(n) = \sum_{i=1}^{k(n)} b_i - a_i$. Then $|M(n) - \lambda S| \leq 2^{-n}$. Hence $\rho_C(\{M(n)\}_{n \in \mathbb{N}}) = x$. Since $\rho_C \equiv \rho_{\mathbb{R}}$, x is computable. ■

Lemma 3.6.5.[11] Let $K \subseteq \mathbb{N}$ and suppose $y = \sum_{n \in K} 2^{-n}$ is a computable number. Then K is recursive.

Proof. Case 1: $y \in \mathbb{D}$ (the dyadic rationals; see Section 3.5). Then y can be written as an infinite binary expansion in exactly two ways: as $(a_1 . a_2 a_3 \ldots a_m 1\ 0\ 0\ 0 \ldots)$ and $(a_1 . a_2 a_3 \ldots a_m 0\ 1\ 1\ 1 \ldots)$, where each $a_i = 0$ or 1. Hence either $n \in K$ for all $n > m$ or $n \notin K$ for all $n > m$. In either case, K is recursive.

Case 2: $y \notin \mathbb{D}$. Then there is a unique binary expansion $(a_1 . a_2 a_3 \ldots)$ of y. By 3.2.7, a number is computable if and only if it is computable with respect to binary notation. Since y is a computable number, the sequence $\{a_i\}$ of its binary digits is

[11] Well known.

recursive. So to determine whether $n \in K$, merely compute the n^{th} digit a_n. Then $n \in K$ iff $a_n = 1$. ■

Proposition 3.6.6 (cf. Ko 1991). There exists a subset of $\mathbb{R}$ that is strongly recursive but not λ-r.a.

Proof.[12] Let K be some non-recursive r.e. subset of $\mathbb{Z}^+$, and let $\{k_i\}_{i \in \mathbb{Z}^+}$ be a recursive enumeration of K without repetition. Construct C just as in the proof of 3.6.3 but with $\{r_i\} = \{2^{-k_i}\}$. Then $\lambda C = 1 - \sum_{n \in K} 2^{-n}$. By Lemma 3.6.5, $\sum_{n \in K} 2^{-n}$ is non-computable, and it follows that $\lambda C = 1 - \sum_{n \in K} 2^{-n}$ is non-computable. Therefore C is not λ-r.a., by 3.6.4 and 3.5.3.

To see that C is strongly recursive, we note that C is closed and show it Δ-decidable. In particular, we give an algorithm to construct a regular Cauchy sequence ϕ for $\Delta_C(x)$, where Δ_C is the symmetric distance function for C. To compute each ϕ_i, proceed as follows: First construct the finite sequence ($a_1 = 0, b_1, a_2, b_2, \ldots, a_k, b_k = 1$) of endpoints of C_i (where $k = 2^i$). Then find some j such that x is in either $(-\infty, 0)$, $(1, \infty)$, $(a_j - 2^{-i}, b_j + 2^{-i})$, or (b_j, a_{j+1}). These sets are r.e. open and cover $\mathbb{R}$, so x can be effectively located within at least one of these sets. If $x \in (-\infty, 0)$, choose ϕ_i such that $|q_{\phi_i} + x| < 2^{-i}$. If $x \in (1, \infty)$, choose ϕ_i such that $|1 + q_{\phi_i} - x| < 2^{-i}$. If $x \in (a_j - 2^{-i}, b_j + 2^{-i})$,

[12] An example of this general kind was suggested by Matthius Schröder in a personal communication.

let $q_{\phi_i} = 0$. Finally, if $x \in (b_j, a_{j+1})$, let $y = \min\{|b_j - x|, |a_{j+1} - x|\}$ and choose ϕ_i such that $|q_{\phi_i} - y| < 2^{-i}$.

Now we merely confirm that $|q_{\phi_i} - \Delta_C(x)| < 2^{-i}$. If $x \in (-\infty, 0)$ or $(1, \infty)$, this is obvious. If $x \in (a_j - 2^{-i}, a_j]$ or $[b_j, b_j + 2^{-i})$, then $q_{\phi_i} = 0$ and $-2^{-i} < \Delta_C(x) < 2^{-i}$, so $|q_{\phi_i} - \Delta_C(x)| < 2^{-i}$. If $x \in (a_j, b_j)$, then either $x \in C$ or $x \in ([a + b]/2 - 2^{-m-1}r_i, [a + b]/2 + 2^{-m-1}r_i)$ for some points $a, b \in C$ and $m > i$. In the former case, $q_{\phi_i} = 0 = \Delta_C(x)$, because C^C is dense. In the latter case, $q_{\phi_i} = 0 < \Delta_C(x) < 2^{-m}r_i = 2^{-m-i-1} < 2^{-i}$, so $|q_{\phi_i} - \Delta_C(x)| < 2^{-i}$. Finally, if $x \in (b_j, a_{j+1})$, then $y = \min\{|b_j - x|, |a_{j+1} - x|\} = \Delta_C(x)$ because $b_j, a_{j+1} \in C$, so again, $|q_{\phi_i} - \Delta_C(x)| < 2^{-i}$. ■

It follows that neither strong recursiveness, Δ-decidability, nor d.i.b. implies either r.a., d.m.z., or the property of being d.i.b. with a measure-zero boundary.

3.7. Conclusions.

Figure 3.3 sums up the relations among the decidability concepts for subsets of $\mathbb{R}^n$ described in this chapter, and two that are not (weak decidability and approximate decidability are defined in Hemmerling 2003). Bold arrows indicate implications first established here or in Parker 2003. Other implications are established in Hemmerling 2003 and Hertling 1999.

A few concepts shown in the graph are accompanied by asterisks. These are concepts for which the logical relations to other concepts in the graph, or the lack thereof, have not been exhaustively established here. Other than those marked with asterisks, properties not connected by a path in the graph are independent. Those independence results not proved explicitly follow from those proved and the implications proved.

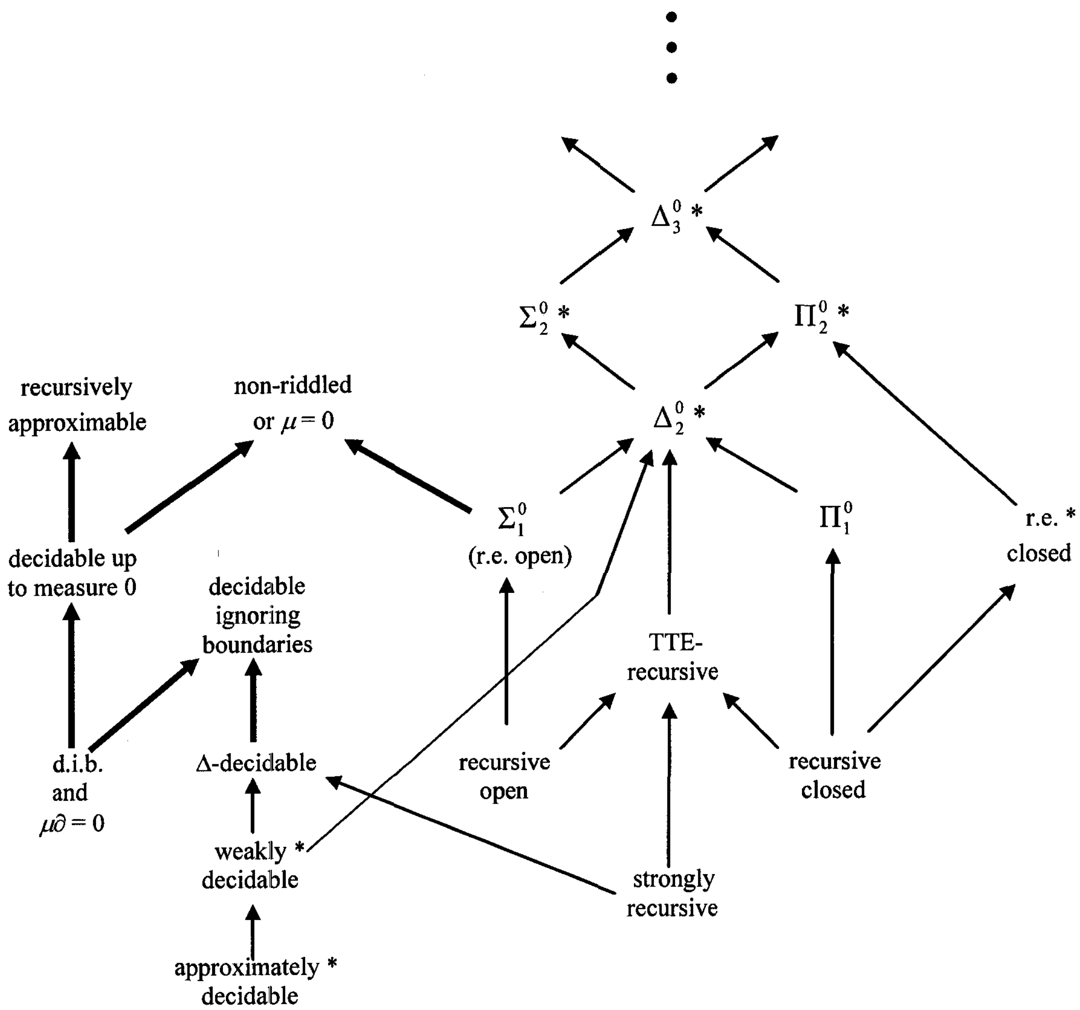

Figure 3.3. Logical implications among notions of a computable set. Asterisks indicate notions for which additional relations of implication have not been strictly ruled out. Bold arrows indicate new results. Also, the implication "decidable up to measure zero ⇒ recursively approximable" shown here is only established in $\mathbb{R}^n$ with Lebesgue measure, though it naturally generalizes to sufficiently well structured measure spaces.

CHAPTER 4

RIDDLED BASINS OF ATTRACTION

4.1. Introduction

Two of the main motivating questions for this dissertation are whether there are actual physical systems that behave non-computably in some meaningful sense, and in *what* sense. In this chapter and the next, we discuss certain mathematical models of physical systems, the indefinite-term (i.t.) behavior of which does exhibit some significant undecidability. We will want to consider what kinds of properties of them are undecidable, what kinds of undecidability those properties have, and how realistic or unrealistic these models are. In the examples at hand, the most significant kind of non-computability that holds, at least among those we have considered in Chapters 2 and 3, is the lack of decidability up to measure zero. Let us use '*non-d.m.z.*' as a name for this property. The cases we consider in this chapter exhibit a particularly strong form of non-d.m.z.; qualitative properties of the i.t. behavior are *nowhere* d.m.z. (This is explained below and formally defined in Definition 3.5.12.) The concepts of d.m.z. and nowhere d.m.z. serve to clarify some non-computability claims that have been made for such systems.

We will need some terminology from the study of dynamical systems. A *dynamical system* is a kind of mathematical model, usually a model of a physical system, but sometimes put to other uses or studied as a purely mathematical object. It consists of a space of possible states—*phase space*—and a function φ called the *flow*, by which the state x at any one instant t_0 determines the state $\varphi(x, t)$ at any other time $t_0 + t$.[1] If φ is defined for $t \in \mathbb{R}$, this is called a *continuous* dynamical system. If φ is defined only for $t \in \mathbb{Z}$ or $t \in \mathbb{N}$, it is a *discrete* dynamical system. The flow of a continuous system is usually determined by differential equations. The phase space often has many dimensions; it enables one to represent all the variables of a system, such as the $6n$ position and momentum coordinates for n bodies, as a single point. The *basin of attraction* (or just *basin*) of a set A of states is the set $\beta(A)$ consisting of those initial states from which the system will asymptotically approach A.[2] For our purposes, an *attractor* is essentially a set whose basin has non-zero volume in phase space. (We will discuss attractors a little more below.)

A basin (or any set) is *riddled,* as in 'riddled with holes,' if every open set in phase space contains a positive-measure portion of the complement of the basin (Alexander et al. 1992; see our Definition 3.6.1). Riddled basins are like Swiss cheese, but Swiss cheese so lacey that no piece, however microscopic, is without holes.[3] Two or

---

[1] This implies that the motion of the system is *autonomous*, i.e., independent of absolute time. However, even a non-autonomous system can be treated as autonomous by introducing an additional state variable that just mirrors the march of time. This also adds a dimension to the phase space. Our main example below from Sommerer and Ott 1996 is just such a case.

[2] More precisely, let X be the phase space. For any $x \in X$ and $A \subseteq X$, let $d(x, A) = \inf_{y \in A} \|x - y\|$, where $\|\cdot\|$ is a norm on X (the standard Euclidean norm if this is well defined). Then for any set $A \subseteq X$, $\beta(A) = \{x \in X: \lim_{t \to \infty} d(\varphi(x, t), A) = 0\}$.

[3] The "holes" here are not necessarily open spaces; they may be disconnected mathematical points or infinitely fine cracks, but arranged so that in any small region of phase space their total volume (measure) is non-zero.

more basins are said to be *intermingled* if they riddle one another, i.e., if every neighborhood containing an element of one basin also contains positive-measure portions of the other(s) (ibid.). Classical models of physical systems have shown numerical evidence of riddled and intermingled basins (Sommerer and Ott 1993, 1996; Heagy, Carroll, and Pecora 1994; Ott et al. 1994, Ott and Sommerer 1994) and some actual systems have shown observational evidence of riddling, or rather approximate[4] riddling (Heagy, Carroll, and Pecora 1994).

A riddled basin implies a kind of unpredictability, since *exact* initial data are required in order to determine whether the state of a system lies in such a basin, and hence to determine the system's qualitative behavior as time increases without bound. This sort of unpredictability is quite different from indeterminism such as that involved in quantum measurement. The models presently in question are completely deterministic: the state at any one instant determines the state at any other. Their unpredictability is also different from so-called *chaos* (i.e., *deterministic* chaos). The central feature of chaos is *sensitive dependence on initial conditions*, the property that for any state there is another arbitrarily close to it in phase space such that the two resulting orbits will eventually separate by some prescribed distance (Devaney 1989). This implies that some *finite*-term predictions require *very precise* knowledge of present conditions. In double contrast, riddled basins imply that some *indefinite*-term predictions require *exact* knowledge of initial conditions (and worse than that: not only do approximate initial

---

[4] The evidence there supports a classical model with riddled basins, but of course such models break down at the quantum level. In that sense the evidence only supports approximate riddling, or perhaps riddling in the classical limit.

conditions fail to determine the i.t. behavior definitively, they fail even to determine it with probability one, as we will see).

Hence sensitive dependence and riddling of basins are seem *prima facie* incomparable; neither is obviously more severe than the other because they concern different prediction tasks. Given some mild conditions, though, riddling is logically stronger: if two spatially separated attractors have riddled basins that are dense in the phase space, sensitive dependence is implied, while on the other hand, a chaotic system may have two separate attractors without riddled basins.[5] Sommerer and Ott (1993) also point out an intuitive sense in which riddling is worse. In merely chaotic systems, it often happens that all orbits are drawn toward one attractor and forever oscillate chaotically near the attractor. Hence all orbits exhibit the same qualitative behavior. Even if there are several distinct attractors, the basins may be solid enough that, except in borderline cases, one can predict which attractor an orbit will approach, given sufficiently precise but inexact initial data. In that case, the general nature of the long-term behavior is quite predictable. But when basins are riddled, we cannot determine which one an orbit lies in, so not only is it difficult to make precise *quantitative* predictions about what state the system will later be in, but even *qualitative* behavior is unpredictable. Without knowing to which attractor if any an orbit is drawn, it might be impossible to tell, for example, what *kind* of oscillations it will undergo or whether it will settle into a particular *region* of phase space. Furthermore, in physical experiments, such qualitative properties of the

[5] The system defined by the so-called two-well Duffing equation is apparently chaotic and has two separate attractors with non-riddled basins (Sommerer and Ott 1996).

behavior could not be consistently reproduced, because we could never prepare the conditions precisely enough to determine the qualitative behavior.

Riddling also implies a kind of *computational* unpredictability. Any computation that determines the indefinite-term behavior of a system whose attractors have riddled basins must actually *use* the complete exact initial data, which in general cannot be finitely expressed. In our view of computation, as conceived by Turing and as actually performed (at least usually), arbitrary real numbers must be represented by infinite strings of symbols (see Sections 2.3 and 3.2.1). Hence a computer (man or machine) cannot make full use of exact real-valued initial data in finite time. Even if exact data are in some sense available, it is impossible to perform a finite computation that depends on them in a critical way.

On essentially this basis, Sommerer and Ott (1996) argue that the behavior of a certain system, which seems to have riddled (and intermingled) basins, is non-computable. However, they do not give a rigorous definition of the non-computability they have in mind. In addition, their results are based on computer-generated images of the basins. As Sommerer and Ott note, one may ask how these computations can be trusted if the basins are indeed non-computable. Sommerer and Ott answer this question, but again in an informal way that we will make precise.

In effect, we have already clarified and bolstered Sommerer and Ott's non-computability claim with Theorem 3.6.2: *no riddled set with positive measure is d.m.z.* Here we will have a look at Sommerer and Ott's dynamical system, as well as a discrete system of simpler construction but with similarly complicated dynamics, to which

Sommerer and Ott draw an analogy. We will see that their arguments and motivations are similar to those behind Theorem 3.6.2 and the notion of d.m.z. We will also note reasons to suspect that for each of the other notions of computable set discussed in Chapters 2 and 3, either Sommerer and Ott's basins are computable, or the notion does not meaningfully apply. In particular, Sommerer and Ott's argument for the validity of their own computations suggests that their basins are recursively approximable (Definition 3.5.1). This explains how they are able to compute reliable graphs of their non-computable basins, as we will see.

However, Sommerer and Ott do not *prove* that their basins are recursively approximable, non-d.m.z., or even riddled, and in this we will not attempt to surpass them. Instead we prove these things (in the appendix) for the simpler discrete system, and we also prove that decidability up to measure zero is the only one of the effectiveness concepts defined in Chapter 3 (except perhaps for the Kleene pointclasses, which we will not consider) that captures the intuitive undecidability of the basins. We will also discuss reasons to expect that those results extend to the more complex and perhaps more physically realistic continuous system.

4.2. Sommerer and Ott's differential equation

Sommerer and Ott's (1996) model describes a point particle moving in a two-dimensional potential, periodically "kicked" by an additional force. The motion is governed by

$$\frac{d^2\mathbf{x}}{dt^2} = -\gamma\frac{d\mathbf{x}}{dt} - \nabla V(\mathbf{x}) + a\sin(\omega t)\,\mathbf{i}, \tag{4.1}$$

where $\mathbf{x} = (x, y)$ varies over $\mathbb{R}^2$, γ is the friction coefficient, $\mathbf{i}$ is the unit vector in the positive x direction, a is the amplitude of the periodic force, $\omega / 2\pi$ is the frequency of the periodic force, and ∇V is the gradient of the potential given by

$$V(\mathbf{x}) = (1 - x^2)^2 + sy^2(x^2 - p) + ky^4. \tag{4.2}$$

The parameters s, p, and k may be varied to obtain a family of potentials. Here we hold all parameters fixed at Sommerer and Ott's chosen values: $\gamma = 0.632$, $a = 1.0688$, $\omega = 2.2136$, $s = 20$, $p = 0.098$, and $k = 10$.

The periodic force $a\sin(\omega t)\,\mathbf{i}$ in Equation 4.1 depends explicitly on time. Consequently, its solutions do not constitute a dynamical system as defined above, for the state at time $t_0 + t$ is not determined by t and the state at time t_0; it depends also on the *value* of t_0—on the time of day, so to speak. However, we can obtain an autonomous (time-independent) system by regarding t itself as a state variable, so that a state for the system is a quintuple $(x, y, dx/dt, dy/dt, t)$. Thus the motion is completely determined by the initial values of the state variables, and solutions form a dynamical system on a phase space with *five* dimensions, corresponding to these five variables.

Intuitively, the system can be visualized as a marble rolling around on the curved surface described by $V(\mathbf{x})$: a deep well with two dips in the bottom, flanking the y-axis, and a small bump between them. (See Figure 4.1.) To incorporate the periodic force,

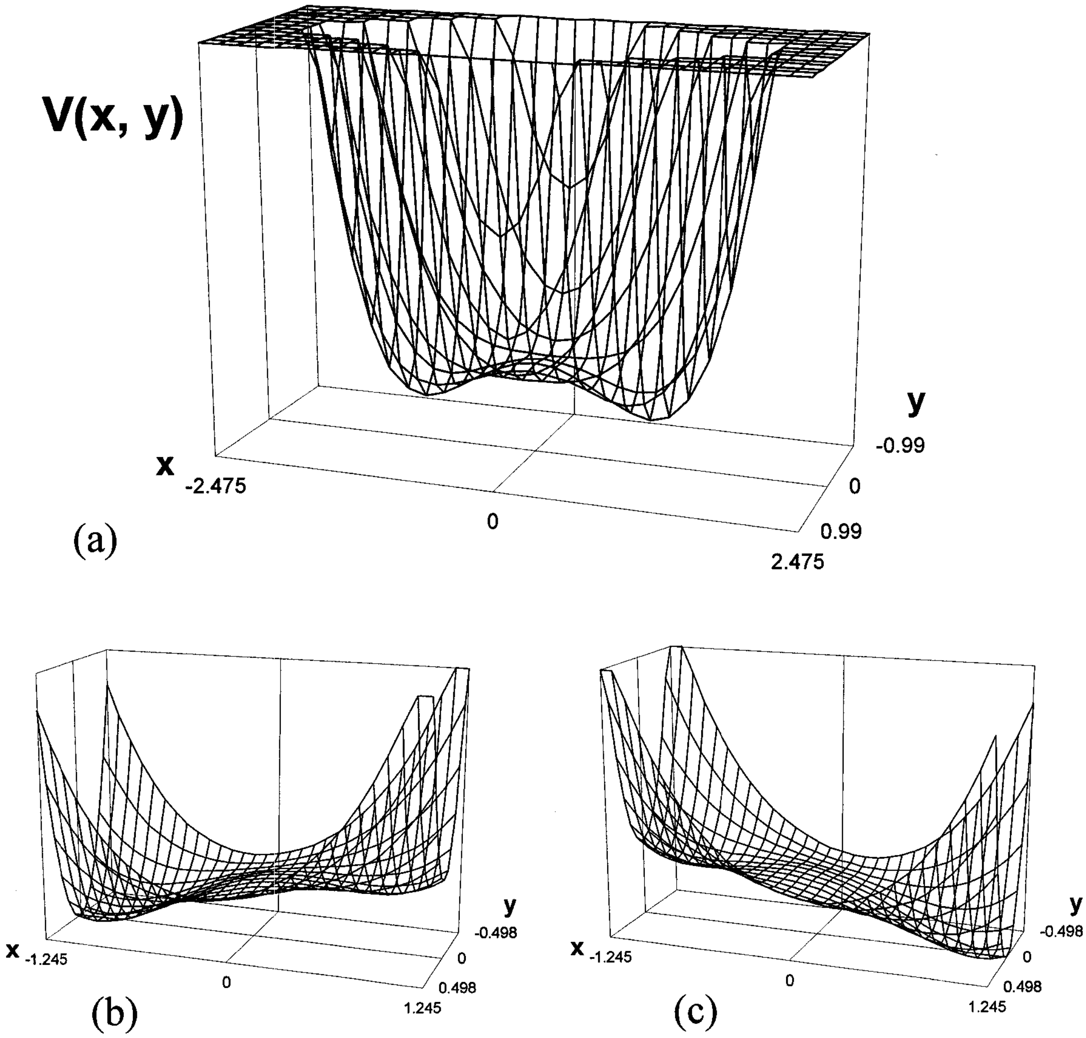

Figure 4.1. Sommerer and Ott's forced potential. The equations describe the motion of a marble rolling on this surface as the surface periodically rocks left and right. Part (a) shows the potential V unaltered (except for the flattening at the top of the graph, which is an artifact of the graphing utility). Parts (b) and (c) show close-ups of the potential with rocking.

imagine this surface rocking gently to the left and right. Due to friction, the marble will tend to settle into one of the two dips, but due to the rocking, it will continue to roll left and right within that dip, near the x-axis. While friction drags the path of the marble down nearer and nearer to the x-axis, where the dip is deepest, the central bump introduces an element of instability; if our marble rolls up onto the bump, it will fall away from the x-axis.

Simulations suggest that this destabilization can occur no matter how closely the marble has settled in near the x-axis (as long as it does not begin *exactly* on the axis with velocity *exactly* parallel to it). Typically, two initial conditions that are very close in phase space result in orbits that remain nearly identical for a brief time, then diverge. (See Figures 4.2 and 4.3.) The position coordinates in the x-y-plane for each orbit soon settle down to a nearly one-dimensional oscillation very near the x-axis, perhaps with both orbits in the left-hand well. Their motion left and right is extremely erratic, driven by the periodic force but not synchronized with it. Occasionally, one orbit may become destabilized, swing around wildly in the left half of the plane, and settle down again. This seems to happen less and less as time passes. Yet at any time, one orbit might become destabilized, swing all the way into the other well, and settle down there, so that the once nearly identical orbits now live in separate dips, probably forever.

Given these seemingly unpredictable episodes of destabilization, one might suppose that orbits never settle down permanently. However, Sommerer and Ott give an analytic (or at least partly analytic) argument that the system does in fact have two attractors corresponding to motion along the x-axis in each of the two dips, so a

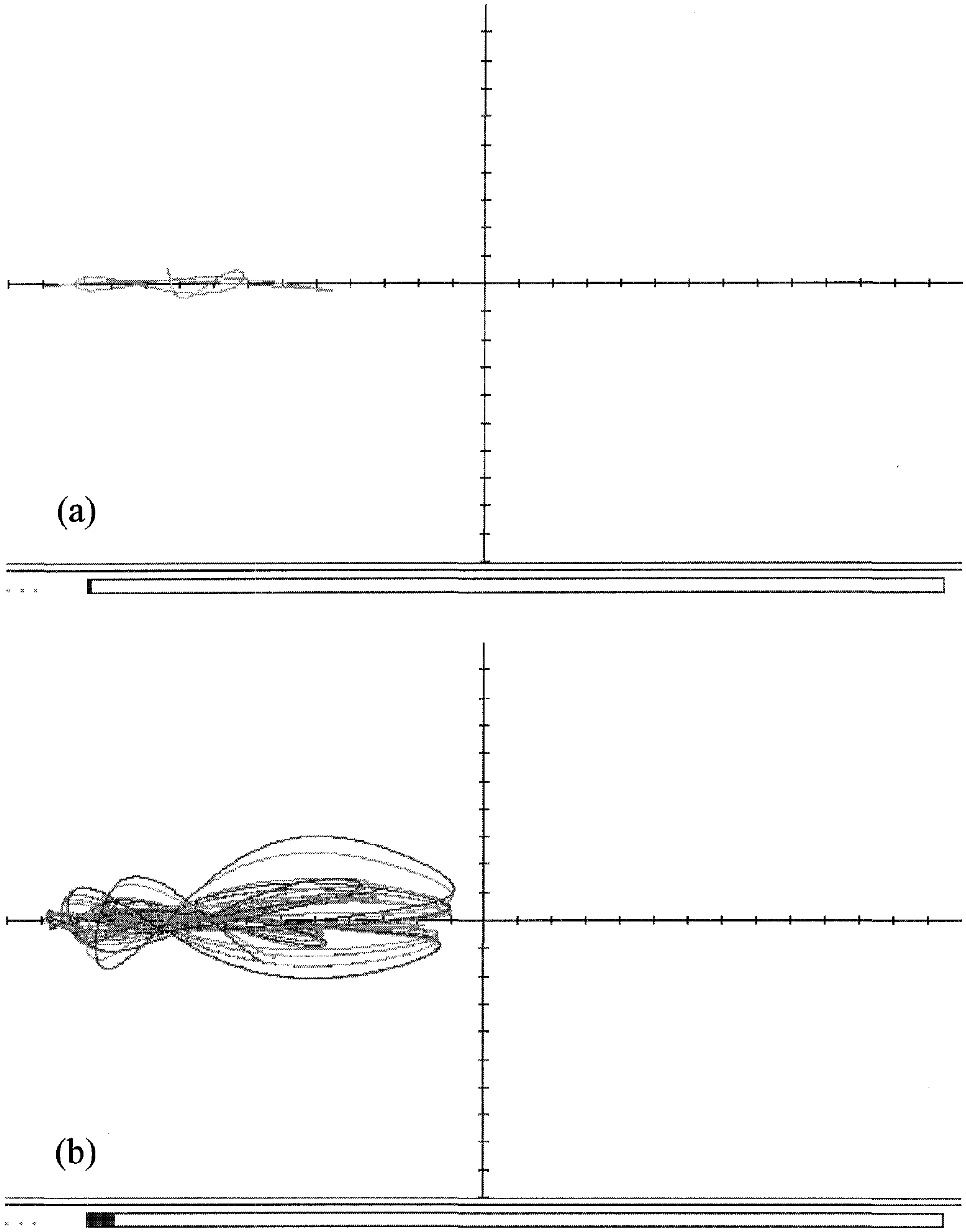

Figure 4.2. Orbits of Sommerer and Ott's system (I). Position coordinates of two orbits with very nearby initial conditions are shown. The length of the red bar at the bottom of each graph indicates time elapsed since the initial states. (a) The two orbits are so similar they are indistinguishable and are shown as one orbit. (b) The orbits soon diverge. (Continued on next page.)

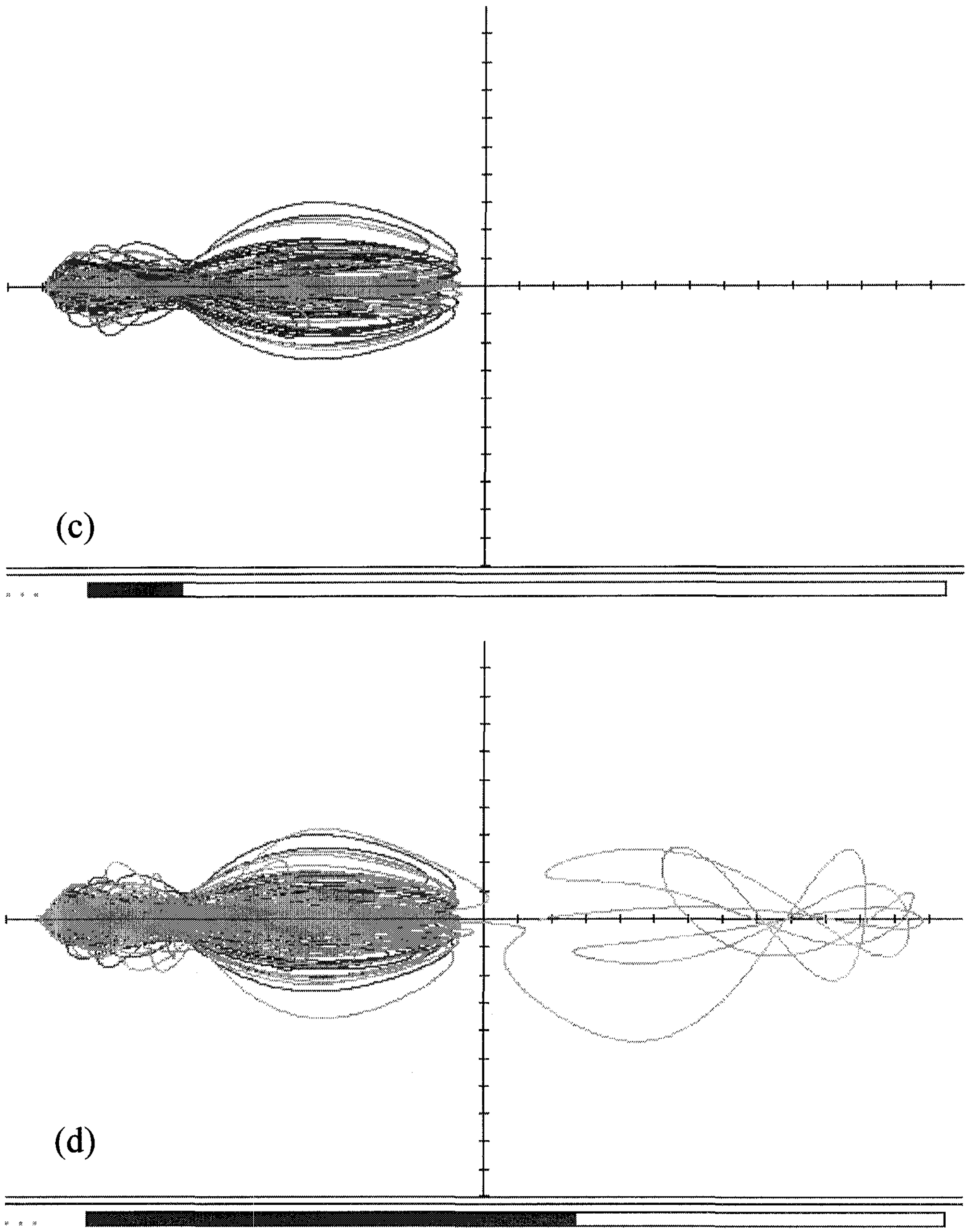

Figure 4.3. Orbits of Sommerer and Ott's system (II). (c) Both orbits soon settle down very close to the x-axis, shown only by the straight red line on the x-axis. (d) After a long time, one orbit becomes erratic and escapes to the other attractor, where it will probably stay. Thus two nearly indistinguishable initial states result in orbits approaching different attractors.

significant portion of orbits do settle down for good. Here, 'attractor' is used in a sense similar to that of Milnor (1985): an attractor is essentially a set whose basin of attraction has positive Lebesgue measure. (Sommerer and Ott also specify that an attractor must be a compact set with a dense orbit. This is slightly different from Milnor's definition, but the differences will not matter for us.) Under other, perhaps more standard definitions of attractor, all points near an attractor must lie in its basin, so an attractor cannot have a riddled basin. Milnor-type attractors are more appropriate to our discussion.

The attractors for Sommerer and Ott's system lie in the three-dimensional subspace where $y = \mathrm{d}y/\mathrm{d}t = 0$, corresponding to marbles that roll along the x-axis exactly. This subspace is an *invariant manifold*: orbits within it stay in it. (This follows immediately from Equations 4.1 and 4.2.) Orbits within this manifold form a dynamical system all their own, governed by the much studied *two-well Duffing equation*,

$$\frac{\mathrm{d}^2 x}{\mathrm{d}t^2} + \gamma \frac{\mathrm{d}x}{\mathrm{d}t} - 4x(1 - x^2) = a\sin(\omega t), \tag{4.3}$$

which we obtain by substituting 0 for both y and $\mathrm{d}y/\mathrm{d}t$ in equations 4.1 and 4.2. The authors make use of two facts about this subsystem that are apparently well known and are confirmed by their simulations (though the present author does not know whether they have been proven analytically): (1) It has two attractors, corresponding to the two dips, *relative to the invariant manifold* (i.e., sets whose basins of attraction have positive *three*-dimensional volume within that manifold). (2) Motion on these attractors is chaotic.[6] Given these assertions, they then argue analytically that the two attractors of the subsystem are also attractors for the larger original system, so many orbits within the five-dimensional phase space are drawn to the attractors within the invariant manifold.[7] That is, in many cases our marble is eventually trapped in one or the other dip. Numerically approximated graphs illustrate the basins of the attractors (Figure 4.4). Both of the basins seem to occupy significant portions of each neighborhood in phase space, suggesting that they are intermingled (and therefore riddled). For the discrete system discussed below, we will prove (in the appendix) that the basins are intermingled.

4.3. Undecidability in Sommerer and Ott's model

Sommerer and Ott infer from their simulations that their basins are riddled and that therefore a computation must make full use of *exact* initial data in order to determine membership in one of these basins. They conclude that the basins are non-computable. They write in their introduction, "[E]ven if the location of an initial condition is available with infinite precision, it is not possible, on the basis of any finite computation, to decide with certainty in which basin the initial condition lies" (244). Yet, this statement fails to specify *for how many initial states* such computation is impossible. If the claim is that no

[6] Sommerer and Ott do not specify the precise sense in which the attractors are chaotic, but they make use of a theorem (due to Alexander et al. 1991) that requires a particular kind of chaos. The conditions of the theorem require that the attractor is ergodic with respect to some invariant probability measure on that attractor. Perhaps this is known to hold for the two-well Duffing equation, but Sommerer and Ott make no case for it in their 1996 paper.

[7] They show that the Lyapunov exponents transverse to the invariant manifold are negative and then appeal to a theorem of Alexander et al. (1991) that if the attractors are chaotic in the sense of footnote 6, these Lyapunov exponents imply that the attractors for the invariant manifold are also attractors for the larger system.

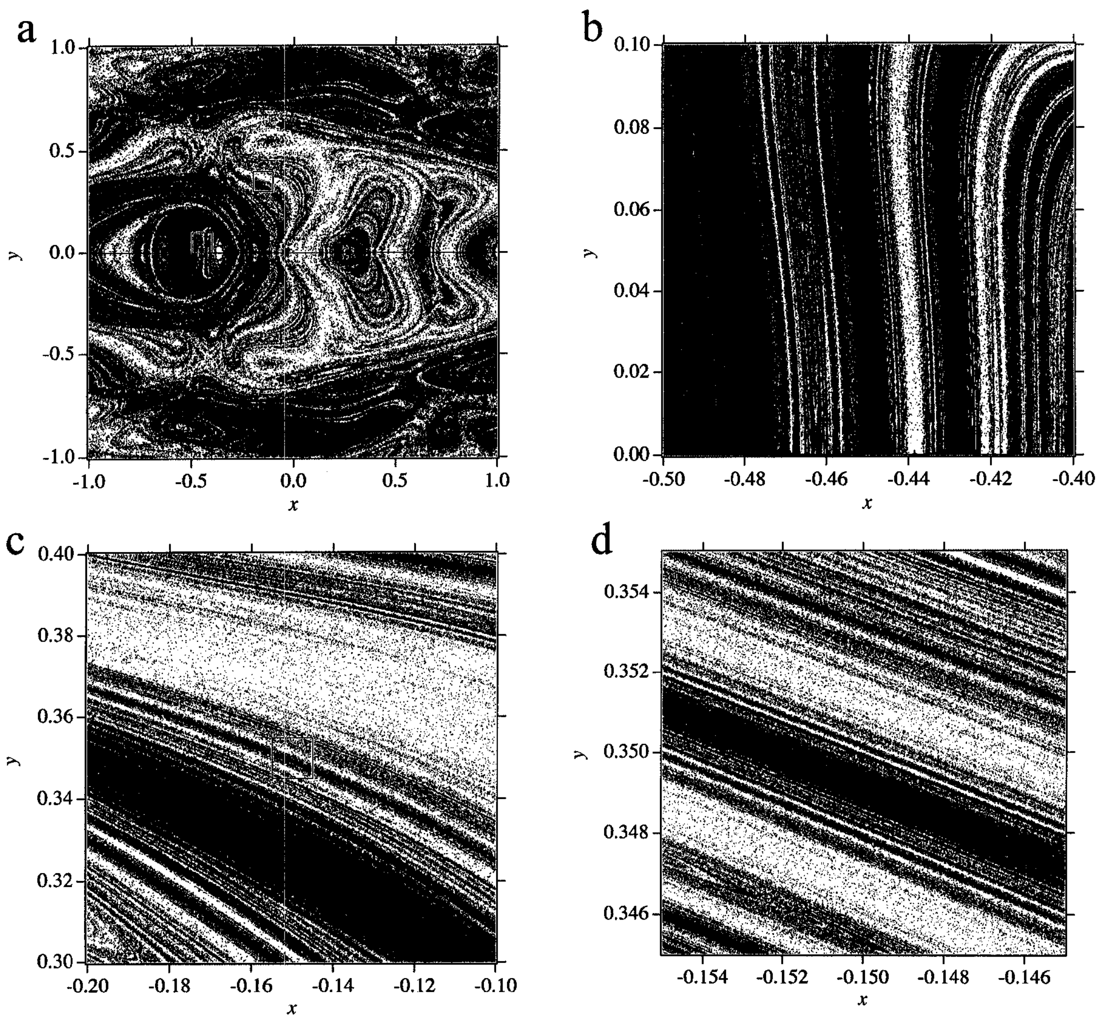

Figure 4.4. Sommerer and Ott's intermingled basins (Sommerer and Ott 1996). Slices of the five-dimensional phase space of the dynamical system defined by Equations 4.1 and 4.2. The attractors intersect these planes along the x-axis. For each initial condition in each 760 × 760 grid, Sommerer and Ott simulated an orbit until it came within 10^{-5} of an attractor, with phase space velocity transverse to the attractor less than 10^{-6}. Initial states leading to the left attractor were colored black, and those leading to the right attractor, white. Blow-ups (b) and (c) of the insets in (a), and blow-up (d) of the inset in (c), suggest that the basins are intermingled.

algorithm can decide which basin *any* state is in, this is false; the algorithm "output yes" will correctly decide *some* initial conditions, unless the basin in question is empty. If the claim is that no single algorithm correctly classifies *all* initial conditions, this is trivial; it is just the lack of what we have called naïve decidability (Definition 3.3.5), a property enjoyed only by the empty set and the entire phase space. Some more specific quantification of the supposed undecidability is required.

What actually makes the undecidability of their example strong is the fact that the basins and their complements both have positive measure in every neighborhood. Therefore, in virtue of the Topological Use Principle (Section 2.4, Proposition 3.2.17), any algorithm will fail to decide membership in the basins correctly not just in a few isolated cases, but on a set of cases with *positive measure*; the basins are not *decidable up to measure zero*. This is our Theorem 3.6.2: No riddled set with positive measure is d.m.z. Note that Sommerer and Ott's intuition that an algorithm cannot actually *use* infinitely precise data about the position of a point is essentially the Topological Use Principle, the main insight used to establish Theorem 3.6.2.

Sommerer and Ott's concern with measure suggests that they had in mind a concept of computability very much like d.m.z., or that they might have come to such a notion if persuaded of the need to clarify theirs. One particular passage further supports this point. Regarding the unavoidable errors in their images of the basins, Sommerer and Ott write,

> [W]e estimate that even for the highest magnification shown in [Figure 4.4], fewer than 1% of the initial conditions are erroneously ascribed to the incorrect attractor. Longer computer runs and greater precision in the computation would

> allow us, at least in principle, to make the error rate arbitrarily small; no *finite* amount of computation, however, could reduce the error rate to zero for *any* system with intermingled basins. (249)

This reference to a percentage suggests a concern with the *measure* of the set of initial conditions incorrectly classified. Their distinction between arbitrarily small error rates and an error rate of zero suggests the distinction between r.a. and d.m.z. Even this quote leaves some room for sharpening, but it at least hints at something *like* a lack of d.m.z., and in any case, non-d.m.z. is one rigorously defined notion of undecidability that does follow from intermingling.

Yet, Sommerer and Ott's basins seem to have an even stronger undecidability. We say a set A is d.m.z. *at a point* x if for some neighborhood U of x, $U \cap A$ is d.m.z. (Definition 3.5.12). Intermingled disjoint sets are not d.m.z. at any points in those sets. Now, Sommerer and Ott assert that their two basins fill the entire phase space up to measure zero (245). This is not obvious; the shape of the potential V makes it clear that there is no attractor at infinity, but this does not rule out the possibility that a non-trivial portion of orbits never permanently settle down to either attractor. Still, it is plausible, and we will show that the basins of an analogous discrete system do fill their phase space up to measure zero. In any case, Sommerer and Ott's simulations at least suggest that every neighborhood of the phase space for their continuous system contains *positive-measure* portions of both basins, and if so, the basins are *nowhere* d.m.z. (This too we will prove for the analogous discrete system.) This means that no algorithm can reliably decide a basin even within a *particular* arbitrarily small neighborhood. We will return to this point in the next chapter.

Most of the other notions of effective set that we have discussed in Chapters 2 and 3 do not meaningfully apply to Sommerer and Ott's basins. Assuming that the basins are indeed intermingled, the TTE notions of strongly recursive, recursive open, recursive closed, r.e. open, Π_1^0, and r.e. closed sets all fail, but for the most trivial reason: the basins are simply not open or closed. They are not open because every open set (apparently) contains portions of the basins' complements, and they are not closed because their complements are not open—every open set contains elements of the basins themselves. The basins are also trivially Δ-decidable and decidable ignoring boundaries. If indeed they permeate every open set, then the symmetric distance function for each basin β is just the constant function $\Delta_\beta(x, y, \mathrm{d}x/\mathrm{d}t, \mathrm{d}y/\mathrm{d}t, t) \equiv 0$, which is of course computable. Further, the interiors and exteriors of the basins are empty, so they are trivially r.e. open, and therefore the basins are d.i.b. To put it another way, the boundary of each basin is everywhere, so ignoring boundaries is ignoring the entire phase space. None of these effectiveness concepts is of much use here.

4.4. Recursive approximability in Sommerer and Ott's model

In all likelihood, though, Sommerer and Ott's basins are also recursively approximable, and this does have some significance. It implies that we can learn *something* about the basins by means of computation, despite their non-computability properties.

Sommerer and Ott provide an informal argument more or less to that effect, though they do not use Ko's term 'recursively approximable' nor establish a similar mathematically precise concept. To generate Figure 4.4, Sommerer and Ott compute, for each of more than two million initial states, an approximation of the resulting orbit, which they extend until the approximated orbit comes very near to one of the attractors.[8] (This can be done with any desired accuracy, thanks to the constructive existence and uniqueness theorem for solutions to differential equations that satisfy the Lipschitz condition; c.f. Earman 1986, 117.) They then color the pixel corresponding to the initial state black or white, depending on which attractor the orbit comes near. Thus they compute images of sets that are supposed to be non-computable! The irony is not lost on them, and they respond as follows:

> One may ask what is the value of computer generated pictures such as those in [Figure 4.4], if proximity to an attractor is not a sufficient condition to determine whether or not the orbit is really destined to limit on the attractor. ...[W]e note that previous results [Ott et al. 1993, 1994] on how the measure of a riddled basin scales with distance from the basin's attractor allow us to estimate the probability of making an error in drawing the computer pictures of [Figure 4.4]. Consider a line segment parallel to, but a distance d away from, the invariant manifold containing the attractor whose basin is riddled. The fraction f of the length of that segment that is *not* in the basin of the attractor scales as $f \sim d^{\eta}$, where $\eta > 0$ is [a constant] given in terms of the (finite-term) Lyapunov exponents. Thus, the closer one gets to the attractor, the greater the probability that one is in its basin. Therefore, when one carries out a numerical simulation, one can quantify the confidence in an initial condition belonging to the riddled basin if its long-time image lies very close to the attractor. ...Longer computer runs and greater precision in the computation would allow us, at least in principle, to make the error rate arbitrarily small... (1996, 248-249)

---

[8] They also require that the orbit's phase space velocity in the directions perpendicular to the invariant manifold is very small before they terminate a simulation. Yet, given the rationale for their technique, which we will discuss, that condition does not seem to be critical.

Thus they argue that there is a procedure to decide membership in the basins with as much statistical accuracy as one likes, *short of* 100 percent. This is the central feature of recursive approximability. They also distinguish (in the previous quotation) between this near-computability and that which their system lacks: the possibility of deciding the basins with *exactly* 100 percent accuracy. (The important distinction they do not mention is that between 100 *percent* accuracy with respect to volume and *perfect* accuracy, point-for-point, which is only possible for the most trivial sets.) Hence r.a. seems to explicate their claims that approximate computation is possible. (This also shows that when they claim that the behavior is *un*decidable, they do not mean in Ko's sense, that of being non-r.a. Another notion is needed to capture their intuition, namely d.m.z.)

The validity argument they sketch above proceeds mainly by analogy, and though it certainly establishes the plausibility of their claim, there are some fine details worth considering. Sommerer and Ott refer to previous results, mainly from Ott et al. 1994, of two kinds: other numerical simulations on other continuous systems similar to the one at hand, which are open to the same questions of validity as the simulations presented in 1996, and analytic results in for discrete-time systems that also exhibit riddled or intermingled basins. Ott et al. (1994) find that for their discrete examples with riddled basins, and apparently (by numerical simulation) for the continuous ones as well, the complement of a basin dwindles near the corresponding attractor in proportion to a power of the distance, as Sommerer and Ott's formula $f \sim d^{\eta}$ indicates. They conjecture that such scaling according to a power law is universal for systems with riddled basins (1994). This is not strictly true, and even in the case of one of their discrete systems it is a kind of simplification (see footnote 10), but it does seem plausible that for any one particular system with riddled basins there will be *some* rule governing the way in which basins scale near the attractors. After all, an orbit is surely more likely to approach an attractor if it is near the attractor than if it is far away, at least in general.

Though, even if this holds for Sommerer and Ott's system, a fully detailed proof that the graphs are largely accurate would involve another a step. Sommerer and Ott's argument suggests that if an orbit comes very near to an attractor A, this is strong evidence that the *initial* condition lies in the basin $\beta(A)$, relative to a probability measure on *all* initial conditions. Yet, even if most states near A lie on orbits that tend toward A in the limit, it could be that a disproportionate measure of initial conditions in the *full* phase space lie on those orbits that come near A but later escape. In that case, the fact that an orbit approaches A would be very little evidence that the initial state from which it was computed lies in $\beta(A)$. To complete the argument, some additional condition is needed.[9] Nonetheless, it is true (and proven) that for the simpler discrete systems of Ott et al. 1994, having an orbit that comes very near to an attractor *is* very strong evidence that the initial condition lies in the corresponding basin, and this does lead us to a proof, quite along Sommerer and Ott's line of thinking, that the basins are r.a. We will discuss this in the next section.

Assuming now that the basins for Equations 4.1 and 4.2 are indeed r.a., this

---

[9] Let $B(A, \varepsilon) = \{x: d(x, A) < \varepsilon\}$ and λ = Lebesgue measure. One condition that would complete the argument is $\lambda\beta(A)/\lambda\{x: (\exists t \in \mathbb{R})\ \varphi(x, t) \in B(A, \varepsilon)\} \approx \lambda[\beta(A) \cap B(A, \varepsilon)]/\lambda\beta(A, \varepsilon)$. In the appendix we prove a similar condition (the Independence Lemma) for the discrete system of Ott et al. 1994, and it seems plausible that some such condition could hold for the continuous system.

provides one clear sense in which graphs of them could be made arbitrarily accurate. As noted in Section 2.7, r.a. implies that there is a method to construct a pixelated graph of the set where the measure of the set of points (not pixels) incorrectly colored is as small as we like. Here we think of each pixel as an open rectangle of exact points, and the error set that we claim can be made arbitrarily small consists of the *points*, within the pixels, that are inappropriately colored. This is apparently not what Sommerer and Ott mean when they claim that the error rate for their graphs is less than one percent. They choose just one point in each pixel and simulate its orbit, so presumably they mean that less than one percent of those *chosen* points are incorrectly colored in their graphs. However, Ott et al. also find that in the discrete systems, the probability that points near to *one another* go to different basins decreases as a power of the distance between them (1994, 392). Hence, even by Sommerer and Ott's graphing method, the color of a pixel might accurately represent the *majority* of points in the pixel, if the pixels are small enough. In any case, if the basins of Sommerer and Ott's system are r.a., as they seem to be, then there is a method by which one *could* in principle effectively generated graphs that are arbitrarily accurate with respect to measure, despite the undecidability of the basins. Thus the distinction between r.a. and d.m.z. accounts in a precise way for the possibility of computing graphs that themselves suggest non-computability.

4.5. The discrete-time system of Ott et al. 1994

Ott et al. discuss two discrete systems in the 1994 paper. We will consider the one studied in their appendix, which is more similar to Sommerer and Ott's continuous system in that it has two bounded attractors with intermingled basins. Sommerer and Ott's argument for the validity of their simulations proceeds largely by analogy with this and similar discrete systems, as does our reasoning that the basins are likely r.a. Therefore we wish to see, in general terms, why the basins of this system are r.a. but not d.m.z., how general this behavior is, and how this system is similar to their continuous system.

We will actually discuss a slight variation on the system of Ott et al. Since we are concerned with the extent to which results about this system generalize to other systems, it behooves us to generalize Ott et al.'s constructions slightly. At the same time, we must impose some minimal computability conditions in order to obtain recursively approximable basins.

Our version of the system consists of iterations of a non-invertible map φ on the rectangle $X = [0, 1] \times [-1, 1]$ with the following general properties:

(i) The effect of φ on y-values is to take them toward 1 or -1, depending on x.

(ii) The effect of φ on x-coordinates is a stretch-and-fold operation similar to the Bernoulli shift map $\phi(x) = 2x \bmod 1$. Hence motion in the x-direction is effectively random, in a sense to be clarified below.

As a result of (i) and (ii), the motion of the y-coordinates will switch directions at random. Nonetheless, we will construct φ so that the upper and lower edges of X, namely $A_{-} = [0, 1] \times \{-1\}$ and $A_{+} = [0, 1] \times \{1\}$, turn out to be attractors.

To do this, we choose a computable function α: $[-1, 1] \to (0, 1)$ defining a curve $x = \alpha(y)$ that divides the rectangle X into left and right sections. (See Figure 4.5.) To

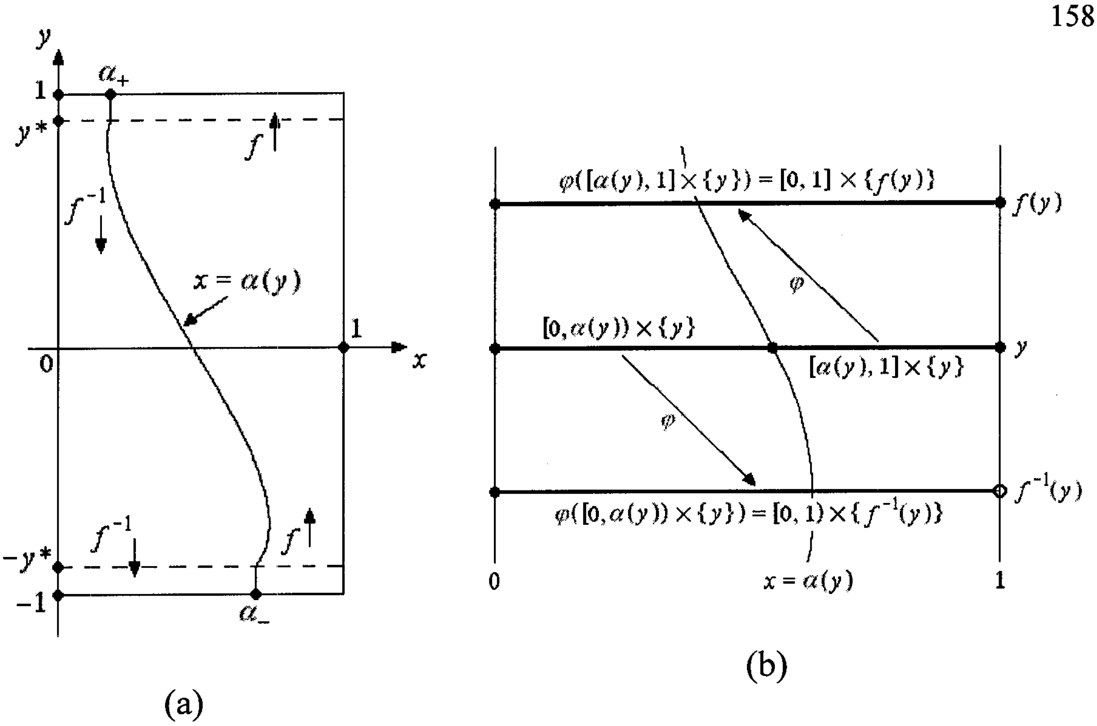

Figure 4.5. A discrete-time dynamical system with intermingled basins. Based on Ott et al. 1994. Part (a) shows the domain of the map φ with schematics of the functions α and f from which φ is constructed. Part (b) shows the operation of φ on a horizontal line segment: the left portion is stretched and lowered while the right portion is stretched and raised.

facilitate the proofs, we assume that for y greater than some y^*, $\alpha(y)$ is a constant α_+ strictly between 0 and 1/2, and for $y < -y^*$, $\alpha(y)$ is a constant α_- strictly between 1/2 and 1. We also choose a computable bijection f: $[-1, 1] \to [-1, 1]$ such that $f^i(y) \to 1$ and $f^{-i}(y) \to -1$ as $i \to \infty$. Now consider a horizontal line segment $[0, 1] \times \{y\}$ across X. Our function φ maps the left portion $[0, \alpha(y)) \times \{y\}$ of that segment downward onto $[0, 1) \times \{f^{-1}(y)\}$, stretching the x-values out linearly by a factor of $1/\alpha(y)$. Similarly, it maps the right-hand portion $[\alpha(y), 1] \times \{y\}$ upward onto $[0, 1] \times \{f(y)\}$, stretching the x-values by a factor of $1/(1 - \alpha(y))$. For a given point (x_0, y_0), we let (x_n, y_n) denote $\varphi^n(x_0, y_0)$. A more formal definition of φ is given in the appendix (Definition A.1).

In Chapter 2, we discussed the probabilistic motivations for our measure-theoretic decidability properties, r.a. and d.m.z. In this section we are interested in whether these properties hold of certain sets, specifically for the standard two-dimensional Lebesgue measure. However, we will first have to consider the one-dimensional Lebesgue measure of certain subsets of horizontal line segments, and it is helpful to regard that measure as a probability. When we refer to the probability that a property P holds of an orbit $\{(x_n, y_n)\}_{n \in \mathbb{N}}$, we mean, for some fixed y_0, the measure of the set of values x_0 such that $P(\{(x_n, y_n)\}_{n \in \mathbb{N}})$ holds (Definition A.4). We are not concerned here whether this notion of probability is physically meaningful; it is just a tool for thinking about Lebesgue measure.

Under iterations of φ, motion in the x-direction is random in the sense that the probable value of x_n is independent of the "coarse-grained history" of the preceding values x_i with respect to the partition imposed by the line $x = \alpha(y)$. More precisely, let H be a set of natural numbers and let H_n be the statement that for all $i < n$, $x_i < \alpha(y_i)$ if and only if $i \in H$ (cf. Definition A.2 in the appendix). Then for any measurable subset S of the interval [0, 1], the probability that x_n is in S is completely independent of H_n; in fact, it is just the measure of S. That is,

$$\Pr(x_n \in S \mid H_n) = \Pr(x_n \in S) = \lambda S.$$

(Cf. Lemma A.5.) Consequently, motion in the y-direction is also random, but in a different sense: the y-values form a *Markov chain*. That is, the probable value of y_n depends only on y_{n-1} and not on any previous values of y; given any n constants y_0', $y_1', \ldots, y_{n-1} \in [-1, 1]$,

$$\Pr(y_n = y_n' \mid y_0 = y_0', y_1 = y_1', \ldots, y_{n-1} = y_{n-1}') = \Pr(y_n = y_n' \mid y_{n-1} = y_{n-1}').$$

The y-motion is equivalent to a "random walk" over the values $f^i(y_0)$, with the probable direction of each n^{th} step determined by $\alpha(y_n)$. (See Lemma A.9.)

Despite this erratic motion in the y-direction, the extreme horizontals A_+ and A_- turn out to be attractors in the long run. This is just because, by our choice of the curve α, orbits near an attractor are more likely to move toward it than away. If an orbit moves even closer to the attractor, it becomes less likely that it will *ever* move away. In fact, all but a measure-zero set of orbits eventually approach one of these attractors (Lemma A.17).

Because α is constant near the attractors, we are able to show in the appendix that (for a given y_0) the probability that an orbit will ever recede from a given attractor decreases in a very regular way as the orbit comes closer to the attractor (Corollary A.11). For example, given that y_n is above y^* by k iterations of f (or more precisely, $f^{k-1}(y^*) \leq y_n < f^k(y^*)$), the probability that some later y_m will ever fall below y^* decreases exponentially with k. Hence the probability that $(x_0, y_0) \in \beta(A_-)$ given that $y_n \geq f^{k-1}(y^*)$ decreases at least as quickly as an exponential function of k. This is a generalization of the key step in Sommerer and Ott's argument for the validity of their graphs,[10] and it is the key step in our proof that the basins are r.a. (Proposition A.21).

That proof is now fairly straightforward. To determine with high confidence in which basin a point lies, we need only approximate its orbit (making use of the fact that f and α are computable) until the orbit comes sufficiently close to one of the attractors. To satisfy the definition of r.a., we must also ensure that this procedure halts on all initial conditions. Notice for example that φ itself is not computable, since it is discontinuous at the curve $x = \alpha(y)$. Attempts to evaluate φ at points on that curve will not halt, so we build into our approximation algorithm a clause that says, if $\varphi^n(x, y)$ ever comes *very* close to that curve, just give up, output zero, and halt. This will result in incorrect outputs, but only for a very small portion of inputs. The full algorithm (excluding details on how to approximate f and α) is given in the proof. The basins are indeed recursively approximable.

Nonetheless, they are not d.m.z., because they are intermingled. No matter how close an orbit comes to an attractor, there is still a small chance it will later defect to the other side. The probability of taking a step away from the attractor may be small, but it

---

[10] Ott et al. (1994, 409) obtain a related result for their particular choice of the function we call f, but their statement appears to contain a typographical error. They use $\Phi_+(z)$ for the probability that an initial state (x_1, y_1) (they use ones instead of zeros here) lies in $\beta(A_-)$ given that y_1 is z steps by f above $y = 0$. They argue that $\Phi_+(z) = A + B(\alpha_+ / \beta_+)$, where A and B are constants and $\beta_+ = 1 - \alpha_+$. They then argue that $A = 0$ and conclude, "$\Phi_+(z) = B(\alpha_+ / \beta_+)^z y_1^{\overline{\eta}_+}$," where $\overline{\eta}_+$ is another constant. Clearly this should read $\Phi_+(z) = B(\alpha_+ / \beta_+)^z$, which by their definitions is equal to $B(1 - y_1 / 1 + y_1)^{\overline{\eta}_+}$. Notice that this is *not* merely a power of $1 - y_1$, and even the fact that it is a kind of power law in terms of y_1 is an artifact of their particular choice of f. Our Corollary A.11 is more general.

has a non-zero minimum, so there is always a non-zero chance of taking many consecutive steps away from the nearest attractor and toward the other. Once near the other attractor, the orbit is more likely to tend toward that one. Therefore every horizontal $[0, 1] \times \{y\}$ contains subsets of both basins, each with positive one-dimensional measure. Now remember what φ does to x-values: it stretches them. Consequently, even a very small horizontal line segment contains a sub-interval that will eventually be stretched all the way across X. Therefore, *every* tiny horizontal line segment contains positive-measure portions of both basins. If we integrate these over any small two-dimensional neighborhood, we find that both basins have positive two-dimensional measure there. Hence the basins are intermingled, and since they permeate every neighborhood, they are not only non-d.m.z. but nowhere d.m.z. (Propositions A.18 and A.19).

These properties are very general. We have made only very broad assumptions about our functions f and α. Even the assumption that α is constant near the attractors can now be relaxed. We have assumed for example that the constant α_+ is between 0 and ½. This enables us to show that the probability that an orbit will ever recede from A_+ decreases in a very regular way near A_+, and this is true even if α_+ is very close to ½. Suppose now that α_+ is indeed very close to 1/2, and suppose we replace the function α with a computable function α' that is not constant near A_+ but is less than α_+ in that region. (See Figure 4.6.) This can only mean that the probability that an orbit will ever recede from A_+ is even *smaller* near A_+. We can still use our algorithm to determine in which basin a point most likely lies, and it will be all the more efficient. Yet, provided

α' has a non-zero minimum, there will still be a positive probability of escaping the attractor, so the basins will still be intermingled and nowhere d.m.z. Consequently, all of our results hold for any computable (and therefore continuous) function α', provided only that $\alpha'(1) < ½ < \alpha'(-1)$, and for all y, $0 < \alpha'(y) < 1$.

We have also assumed that, in a single iteration, φ stretches both the left and right portion of a horizontal all the way across the domain X. This too can be relaxed. As long as φ affects some minimal stretching, so that every small line segment is eventually stretched enough to include portions of both basins, we still have intermingled basins. In fact, the particular form of the motion in the x direction does not matter much, so long as

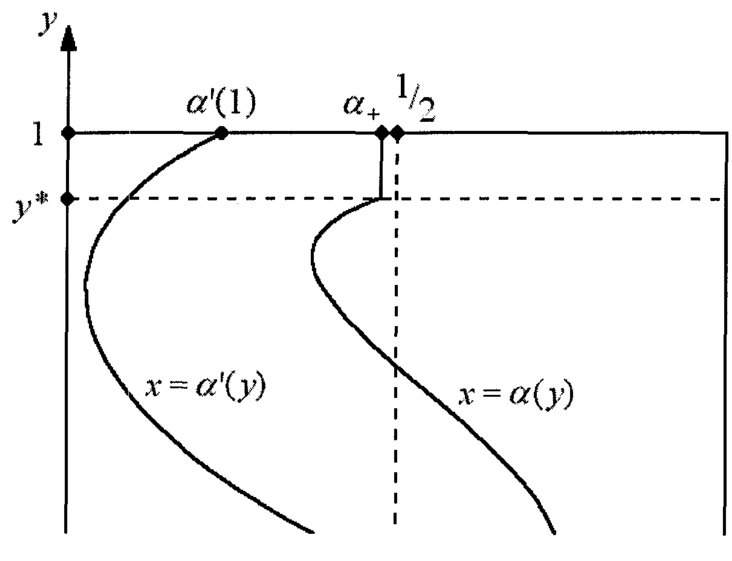

Figure 4.6. Generalizing the discrete system. If we replace α with another computable function α': $[-1, 1] \to (0, 1)$ such that $\alpha' < \alpha_+$ for $y > y^*$, the basins are still intermingled, non-d.m.z., and recursively approximable. This holds even for α_+ very close to 1/2 and y^* very close to 1, so α' may be any computable function with $\alpha'(1) <$ 1/2 and (by a parallel argument) $1/2 < \alpha'(-1)$.

almost all orbits spend some time in regions that are drawn toward a given attractor (such as $x < \alpha(y)$) and some in regions that are pushed away ($x > \alpha(y)$).

In this respect, Sommerer and Ott's continuous system is very much like the discrete system. The two attractors have stable regions (the dips in the potential) and unstable regions (the central hump). Motion near the attractors is chaotic, spending some time in both the stable and unstable regions. The similarity between the two systems becomes more striking if we just modify the discrete system a little (Figure 4.7): translate and bend the rectangle X so that both attractors lie on the x-axis, with the unstable regions close together. Then produce the mirror image of this figure below the x-axis. As in the continuous system (looking only at the x- and y-dimensions), we now have two chaotic attractors on the x axis that are largely stable, but unstable in a small central region. It seems very plausible that, like those of the discrete system, the basins of the continuous system are indeed intermingled but r.a. Ott et al. may have over-stated their conjecture that riddled basins always scale according to a power law near attractors; in the case of the discrete map φ, this depends on the particular choice of f. Yet, it also seems likely that for a continuous system like theirs there will be *some* rule to the effect that orbits near an attractor tend toward it, so that, if the differential equations are sufficiently regular (i.e., Lipschitz) and computable, the basins will be r.a.

4.6. The physicality of Sommerer and Ott's model

Sommerer and Ott complain that Cristopher Moore's dynamical systems (which also are claimed to exhibit non-computable indefinite-term behavior; 1990, 1991) are not

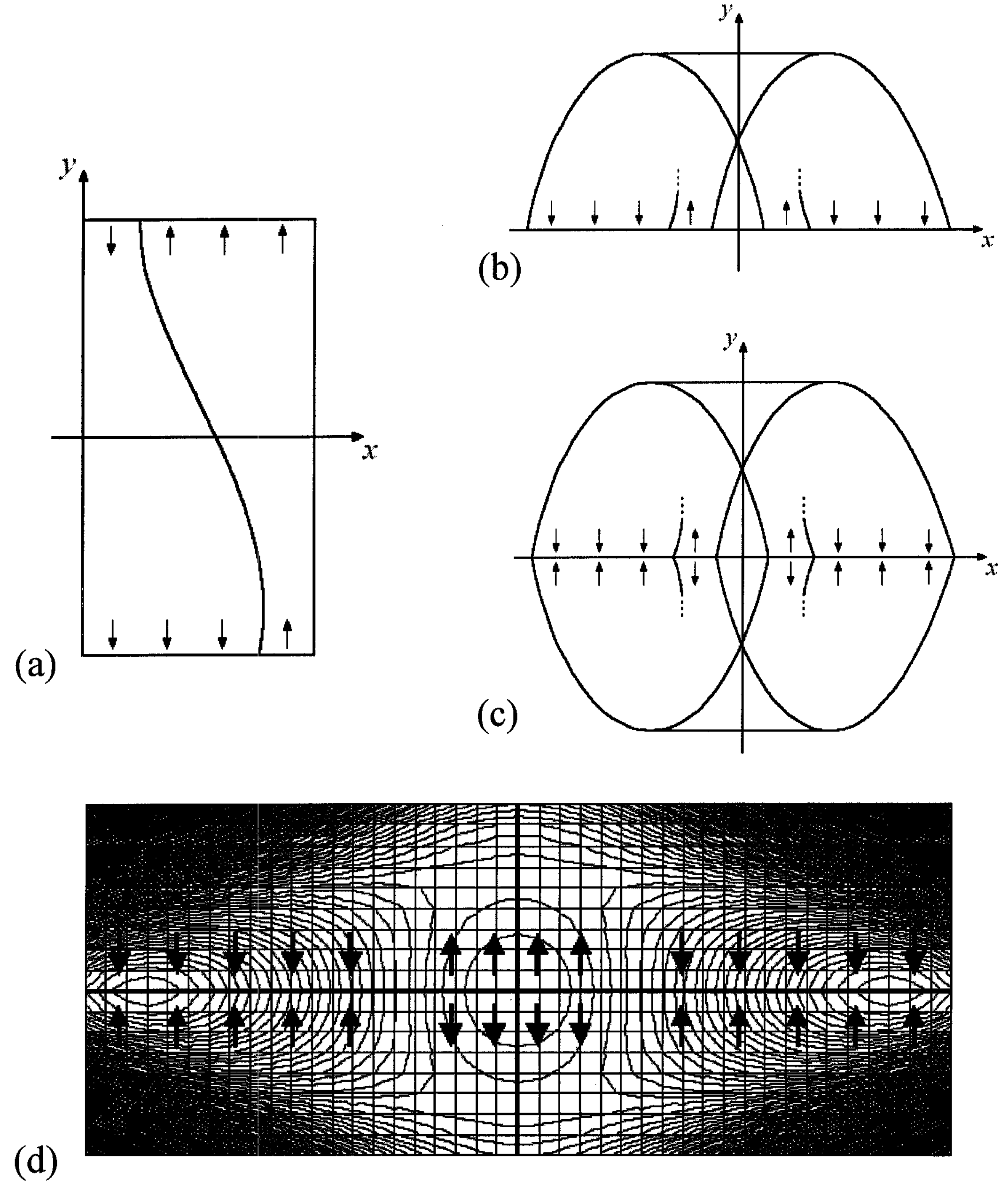

Figure 4.7. Comparison of the discrete and continuous systems. (a) The original domain X of the discrete system, with arrows showing attracting and repelling regions of the chaotic attractors. (b) X translated and folded. (c) A reflection of the result is appended to the system to make the similarity to the continuous system apparent. (d) A contour graph of the potential V of the continuous system, showing attracting and repelling regions of the x-axis, where the two chaotic attractors intersect this graph.

very "physical," i.e., not realistic models of physical systems. But is Sommerer and Ott's system very physical? Might this kind of unpredictability exist in real physical systems? There are at least three issues in particular to consider here: (1) Are there likely to exist systems with roughly the same dynamics? (2) How does noise affect the dynamics? And (3) do not the model and its computability and non-computability properties break down at the quantum level?

It is hardly commonplace to stumble on a surface in the shape of the potential V steadily tilting left and right with a marble on it. However, it would not be impossible to manufacture such a system, and accurately enough that its behavior would reflect the riddled structure of Sommerer and Ott's basins. What makes this plausible is the fact that the riddling is *structurally stable*: it survives small changes in the parameters of the motion. The relevant variables include the amount of friction, the frequency and amplitude of the periodic force, and perhaps most significantly, the shape of the surface, or the potential V. If we were to manufacture a rocking two-welled dish in hopes of reproducing the erratic behavior of Sommerer and Ott's model, the shape of the dish would not have to be exact . The same qualitative dynamics persist over a range of different potentials defined by varying the parameters. Ott et al. (1994) explicitly report this for a potential qualitatively different from V, and Kan (1993) *proves* an analogous result for certain discrete-time systems on the thickened torus $T^2 \times [0, 1]$.[11] We have also

[11] Specifically he shows that for some k there is an open set of C^k diffeomorphisms of the thickened torus having two attractors with intermingled basins (that is, an open set *among* the diffeomorphisms of the thickened torus, all of which preserve the invariant manifolds $T^2 \times \{0\}$ and $T^2 \times \{1\}$).

seen in intuitive terms that the riddled basins in the discrete system of Ott et al. 1994 persist under many broad variations. Sommerer and Ott (1996) at least suggest that the same holds for their 1996 continuos-time system, and a handful of simulations conducted by the present author using the "Phaser" software tend to confirm this.

Intuitively, some insensitivity to the shape of the potential is to be expected, since, as we have noted, the riddling results from a few of its gross features: the two dips, which tend to draw orbits in if friction is present, and the central hump, which tends to destabilize orbits that run near the invariant manifold. Sommerer and Ott emphasize the importance of maintaining the y-symmetry in Equations 4.1 and 4.2 in order to maintain riddling, and of similar symmetries in other riddled systems. The symmetry guarantees the existence of the invariant manifold where the attractors are found, a critical part of the analysis by Sommerer and Ott and by Alexander et al. (1992). However, even this symmetry is not strictly necessary. One can always make a change of coordinates that destroys the y-symmetry without changing the dynamics, or to put it another way, one can apply nearly any homeomorphism to the system and just let the new dynamics be defined by applying the same transformation to the old dynamics. The new basins would just be the images of the old basins, and if the homeomorphism preserves sets of positive measure, riddling is maintained. So the important thing is not the symmetry *per se* but having a chaotic attractor in an invariant manifold. Yet, perhaps even an attractor that is not strictly contained in an invariant manifold could also generate riddling. That remains to be seen.

The rocking of our fabricated surface would not have to be exactly as prescribed by Sommerer and Ott's equation either, in order to generate riddled basins. The important thing there is that it should keep the marble moving left and right chaotically. Just about any oscillating force with *approximately* the right direction, amplitude, and frequency would surely suffice.

Of course, our real interest is not in marbles on strange surfaces, but in systems in general with qualitatively similar behavior. Heagy, Carroll, and Pecora (1994) have observed an actual electrical circuit, the behavior of which shows clear evidence of a riddled basin with positive measure, or at least an approximately riddled basin. That basin is apparently *somewhere* d.m.z. (Definition 3.5.12), but not d.m.z. on the whole, at least insofar as it really is riddled. Ott and Sommerer (1994) suggest that riddled basins might also occur in some chemical reaction-diffusion systems.

Noise is another real-world factor that undermines models like Sommerer and Ott's, and this does have a major impact on the dynamics. Ott et al. suggest that low noise would have the effect of disturbing even those orbits that would otherwise have settled down to an attractor (1994, 392). There would still be an *appearance* of riddled basins, but they would not really be basins of attraction as almost no orbits would approach an attractor in the limit. Rather, almost all orbits would behave in the same general way, forever jumping from one near-attractor to the other, so the qualitative behavior would become trivially d.m.z. Heagy et al. observe in their circuit that the power-law scaling of the (apparent) basin levels off very near the (apparent) attractor, and they attribute this to noise. True riddling, it seems, is reserved for truly isolated, deterministic systems.

Further, these models and their basin structure seem to break down at the quantum level. A basin is only really riddled if even the smallest neighborhoods of phase space contain positive-measure portions of its complement. Computability, and in particular d.m.z., is also an absolute, ideal concept, and whether or not a set is d.m.z. depends on arbitrarily small details of the set's structure. If riddling is to be the basis of non-d.m.z. character, there must be no scale at which the riddling ceases. For extremely tiny differences in initial conditions, quantum mechanics predicts behavior quite different from that observed at the macroscopic level, and even the general framework of precise, unique positions and velocities represented in a finite-dimensional phase space becomes inadequate. Whether quantum mechanics permits any undecidability related to non-d.m.z. lies a little outside the scope of this project, but we will touch on it again in Chapter 6, where we consider more deeply the question of undecidability in real systems.

Regardless of this question of physicality, the concept of d.m.z., or non-d.m.z., captures a kind of undecidability manifested in some natural physical models of a very conventional kind—systems of smooth ordinary differential equations. Furthermore, it clarifies just what kind of undecidability systems like Sommerer and Ott's—systems with riddled basins—have. D.m.z. is the only one of the many decidability concepts we have reviewed that fails for their examples in a non-trivial way. The argument that Sommerer and Ott's basins are non-d.m.z. reflects those authors' own reasons for claiming non-computability, while it also distinguishes that non-computability from both the trivial "naïve" non-computability common to almost all sets and from the non-trivial computability that their basins do seem to have: recursive approximability.

CHAPTER 5

FURTHER RESEARCH AND CONCLUSIONS: THE KAM TORI, THE STABILITY OF THE SOLAR SYSTEM, AND UNDECIDABILITY IN REAL PHYSICAL SYSTEMS

We live in an old chaos of the sun...

Wallace Stevens, "Sunday Morning"

5.1. Introduction

Some further extensions of the research presented in the preceding chapters have already been carried out but are not yet ready for presentation in their final form. These include applications of decidability concepts and theorems to nearly integrable Hamiltonian mechanical systems, with possible implications for the mathematical problem of the stability of the solar system; a proposed way of extending decidability properties from mathematical sets to actual physical systems; and an examination of the significance of undecidability in classical models for our understanding of the actual world. Some of the results of that research are summarized here.

5.2. The KAM tori and the stability of the solar system

The most noteworthy application so far of the concept of decidability up to measure zero is to the infamous problem of the stability of the solar system discussed in

Chapter 1. As noted there, some have suggested that this problem may be undecidable (Moser 1978; Wolfram 1985, 2002), but without articulating an appropriate non-trivial meaning for this claim.[1] The concept of decidability up to measure zero fills this gap. In fact, KAM theory (named for its originators, Kolmogorov, Arnol'd, and Moser), provides strong reasons to suspect that the stability of planetary systems and many related problems are not decidable up to measure zero. It shows that for certain dynamical systems ("nearly integrable Hamiltonian systems"), including idealized planetary systems, many possible initial conditions result in bounded orbits confined to tori in phase space.[2] For a single system of the right kind, these tori form what Moser calls "a set of positive measure...but a complicated Cantor set" (1973, 8). Like the generalized Cantor set that we saw as an example of a non-d.m.z. set in Chapter 3 (Theorem 3.6.3), the union of the KAM tori has positive measure, is riddled, and is therefore not decidable up to measure zero. Any algorithm to decide whether a given point lies on one of the KAM tori will fail in a non-zero percentage of cases.

The fact that the KAM tori for a given system are not d.m.z. does not directly imply that the stability of planetary systems is similarly undecidable. This depends on

---

[1] Arnol'd does articulate a more concrete decidability question, but of a rather different kind from that considered here and for a different class of stability problems (Browder 1976, 59). He develops this further in a letter published in da Costa and Doria 1993.

[2] A torus is essentially a doughnut surface, but possibly of higher dimensions. It can be constructed by identifying, for each coordinate, any two values that differ by an integer multiple of some particular, specified, positive real number k, so that $(x_1,\ldots, x_m,\ldots, x_n) \equiv (x_1,\ldots, x_m + k,\ldots, x_n)$ for any $m \in \{1,\ldots, n\}$ and any $x_1,\ldots, x_n$. Note that after these "identifications" are made, addition and subtraction of coordinate values still make sense (since if a and a' differ by a multiple of k, and if b and b' also differ by a multiple of k, then so do $a + b$ and $a' + b'$ as well as $a - b$ and $a' - b'$. An n-tuple of angle measures ranges over a torus.

the behavior of those orbits that do not lie on the KAM tori, about which little is known. However, it is known that some escape (Arnol'd 1966), and it would be quite surprising if each of these did not take with it some open set of nearby orbits.[3] If indeed the full set of stable orbits for a given planetary system has positive measure (as we know it does in some cases) and is riddled with escaping orbits, then it is not d.m.z. In that case we can conclude that the stability of planetary systems is undecidable in this precise sense.

The KAM tori for a system are computable in most if not all other senses, at least if all of the parameters and coefficients in the equations governing the system are computable. The tori are explicitly constructed by an iterative process (Arnol'd 1963) of removing open sets, much like that by which a Cantor set is constructed (see the proof or Proposition 3.6.3). As in the latter case, this iterative process itself provides a means of deciding membership in the KAM tori up to an error set with arbitrarily small measure, witnessing that the union of the tori is recursively approximable. The gaps between them are r.e. open, and their interior is empty, so they are d.i.b. They form a closed set, and the distance to the nearest KAM torus is computable, so they are recursive closed and even strongly recursive. All of this holds, that is, provided the parameters and coefficients involved in the construction of the tori are computable. In the case of a planetary system, for example, the relevant parameters are the masses of the bodies and the gravitational constant. However, even if these quantities are not computable numbers, the union of the

---

[3] The suggestion intended here is not that the set of escaping orbits is open but that it is a *regular* closed set, i.e., the closure of its own interior. In the two-body case, for example, there are bounded elliptical orbits (elliptical in the space of *positions*, that is), escaping hyperbolic orbits, and "between" these (in phase space), a boundary of parabolic orbits that also escape. The phase points corresponding to the elliptical and parabolic orbits form a regular closed set.

KAM tori is at least r.a., d.i.b., and strongly recursive *relative* to such parameters. That is, there exist algorithms that satisfy the definitions of those computability concepts *if* provided access to the parameters for reference, in addition to the initial conditions.

Whether or not the stability of planetary systems is ultimately d.m.z., the fact that the *KAM tori* for a given system are not d.m.z. is a meaningful undecidability result for an important area of mathematics, and along with the positive decidability properties of the tori, it clarifies the precise senses in which stability problems may be rigorously unsolvable. It also illustrates the usefulness of d.m.z. in bringing undecidability to light in cases where the other concepts do not. The fact that any proposed decision procedure for the KAM tori of a system will fail on a positive-measure set of cases is an insight that none of the other concepts of computable set facilitates.

5.3. Decidability for real physical systems

Some authors seem to suggest not only that certain classes of *mathematical models* have undecidable indefinite-term behavior but that the behavior of certain actual *physical* systems might be undecidable. We have made some progress in understanding what it could mean for a continuous mathematical model to exhibit undecidable behavior, but can we somehow apply these concepts to actual systems? After all, decidability as we have understood it so far is a property of *sets*—abstract mathematical objects—whether sets of formulas, integers, or (for appropriately relaxed concepts of decidability like d.m.z.) points in a continuous space. Can we make sense of the idea that some such decidability or undecidability lies in real systems out there, and is not just an artifact of our favorite model? Also, the claim that there is *actual* undecidability and consequent unpredictability in the world seems to assert that some *individual* system in a particular state is computationally unpredictable. Usually, undecidability means that there is no one algorithm to decide whether an object is in a given set in *all* or (for non-d.m.z. or non-r.a. sets) *most* possible cases. (The "possible cases" would be natural numbers, points in a space, or in general, the objects in some superset of the set in question.) Can we somehow make sense of undecidability for a *single* system, where there is only one initial state to be classified? Finally, even given all of this, what can we make of the undecidability results we have obtained? Is it not the case that all of our models are inaccurate anyway? And is not undecidability in classical systems immaterial if the world is ultimately non-deterministic at the level of quantum mechanics?

5.3.1. Mathematical properties of real systems

As Myrvold writes, "It is, after all, theories, not things, which are formulated mathematically, and to which mathematical notions such as computability and constructivity apply" (1995, 33-34). ("Theories" here refers not to sets of sentences in formal languages, but to mathematical models like dynamical systems.) This is strictly true, but there is a way in which mathematical properties might reasonably be attributed to real physical systems, provided a certain amount of physical realism. The idea applies more precisely to *collections of possible states* for *classes of similar systems*—similar, that is, in their dynamics but varying in their initial conditions and perhaps in *some* features of their dynamics, corresponding to parameters.

Suppose we are concerned with a particular class of systems, such as systems of three nearly spherical bodies in otherwise nearly empty space. Suppose that such systems have certain physical properties—their "states"—that are *naturally* correlated with points in $\mathbb{R}^n$ by means of some physical measurement procedure. Suppose also that there are facts about how such systems *would* behave if they were in various states. Then we could reasonably say that a collection of states "has" a mathematical property P if the corresponding set of points in $\mathbb{R}^n$ has property P. Thus we might say, for example, that stability is non-d.m.z. for physical systems of three bodies (not just models) if the set of those states that *would* lead to escape or collision corresponds to a set in $\mathbb{R}^n$ that is not d.m.z.

Attaching properties to a class of physical systems in this way does not differ substantially from attaching them only to a mathematical model, if that model is sufficiently accurate. However, two features of the proposal justify regarding the properties as intrinsic to the physical systems themselves: (1) The mathematization of the states of physical systems is (by assumption) *natural*, given by the physical results of physical measurement procedures. Hence a property of the mathematical points corresponding to a collection of physical states is not an artifact of an arbitrary association but a physical fact about the outcomes that would result from certain physical procedures. And, (2) the collections of states to which we propose to attribute such properties are not determined by our currently favored dynamical theory, but by physical facts about how such systems really would behave given various initial conditions. Hence if a property is attributed to a collection of states in this way, it is not in virtue of any arbitrary theoretical choices, but only in virtue of physical counterfactual truths.

This analysis of course involves complications to address, such as the possibility that different measurement procedures might result in different attributions of properties. One might also raise generally anti-realist or skeptical objections to the notion of physical states or the hypothesis of counterfactual truths on which this account depends. Rather than enter into such issues, we let the analysis (here brief and sketchy) stand as a kind of framework for associating mathematical properties such as undecidability with whatever real systems there might be. We merely propose a manner of speaking, not a thesis about reality. However, remarks like Myrvold's might have led one to believe that even assuming realism, one could not meaningfully speak in such a way.

5.3.2. Classes of systems and individual systems

Does it make sense to suppose that problems like that of the stability of *the* solar system, of our actual system in its particular state, are somehow undecidable? Perhaps we can at least *make* some sense of this. If we have a decision procedure that works for almost all states corresponding to a particular region of phase space, and the state of the actual system happens to correspond to a point in that region, then we can say that the property in question is locally d.m.z. for that particular system. Intuitively, this implies that in that particular case we can decide with probability one whether the property holds. If the structure of the KAM tori is any clue, it seems that the stability of a given many-body planetary system is locally d.m.z. if and only if the system is unstable. In contrast, Sommerer and Ott's basins of attraction (1996) are not locally d.m.z. at *any* point. Even within any tiny neighborhood, no algorithm can decide their membership with probability one.

5.4. The inexactness of models

Do we have any reason to suspect that undecidable behavior of some sort does occur in actual physical systems? The models in which we have discovered undecidable behavior are, after all, highly idealized, and d.m.z. is a very subtle property, depending on the finest details of sets.

However, many of the details of real-world systems omitted by our models may amount to small perturbations, and the undecidability we have seen might survive such disturbances. We noted in Chapter 4 that the riddling of basins in Sommerer and Ott's systems (1996) survives some variations. The structure of the KAM tori for a nearly integrable Hamiltonian system also survives some perturbations; it is part of the explicit content of KAM theorems that the tori persist to some extent as certain parameters are varied (e.g., Arnol'd's μ, a factor in the masses of the smaller satellites of a planetary system; 1963). On the other hand, friction or any kind of dissipation immediately destroys the Cantor-like structure of the tori (Feudel and Grebogi 1997; Pöschel 2001). However, some invariant tori survive and become attractors, and their basins have a very complicated structure (Feudel and Grebogi 1997). Perhaps these basins are non-d.m.z. This remains to be explored.

As previously remarked, these classical models ultimately break down at the quantum scale. If one takes a reductionist view of physics, any undecidability of the kind we have discussed in actual systems must be grounded not in classical models but in the most microscopic features of the world. One might argue, against such reductionism, that classical and semi-classical descriptions of the large- or intermediate-scale features of systems have a legitimacy, explanatory indispensability, and verity of their own. However, even if not, there may be interesting decision problems within the quantum formalism. As mentioned in footnote 27 of Chapter 2, Wayne Myrvold (1997) has considered whether entanglement is a decidable property of quantum states (a problem, he reports, that came to Abner Shimony in a dream). Myrvold shows that the set of entangled states in a product space is dense, r.e. open, d.i.b., and (in effect) strongly recursive.[4] To determine whether this or any other interesting set of quantum states is d.m.z. will require some further thought about the appropriate measure to apply, and for that matter, the appropriate topology.

Though the question is far from settled, these considerations suggest that it is at least conceivable that some actual physical systems exhibit some genuine undecidability. It is not *prima facie* out of the question, and we now have a clearer idea of what we might *mean* by proposing it.

---

[4] Myrvold does not employ the notion of strong recursiveness, but he shows that the set of entangled states is r.e. open, the distance to its complement is a computable function, and its exterior is empty. This implies strong recursiveness.

5.5. Final conclusions

It has been a very long journey indeed, and we have seen much along the way. Let us review.

We set out to understand better the limitations of any systematic reasoning or procedures as applied to the physical world. We stated as a motivating question, "What can and cannot we learn about the behavior of a physical system by systematic calculation on the basis of a real-valued model?" Our main concrete goal was to see, both in the sense of argument and that of illustration, that decidability up to measure zero is an especially appropriate concept of decidability for application to problems in physics, and in particular that it is the most significant kind of decidability implied by arguments like Sommerer and Ott's, based on the riddled structure of sets.

We considered at some length the value of investigating such undecidability in the indefinite-term behavior of systems. We compared and contrasted different models of computation and different concepts of recursive and decidable sets in $\mathbb{R}^n$. We saw that, unlike the other notions, d.m.z. is strong and strict like classical discrete decidability, appropriately symmetric with respect to sets and their complements, and well motivated by probabilistic physical considerations. We saw that behavior of Sommerer and Ott's (1996) dynamical system is likely not d.m.z., though it is r.a., enabling one to study it by means of computations. We stated, without much explanation, that the union of the invariant KAM tori for a nearly integrable Hamiltonian system is also riddled and has positive measure. Therefore, such a set is also non-d.m.z., and this undecidability may well extend to the stability problem for idealized planetary systems (and other systems), even while all other kinds of decidability hold. Finally, we briefly suggested ways to apply our mathematical decidability concepts to actual physical systems, and the prospects for finding decidability there. We have concluded, if tentatively, that undecidable behavior in actual physical systems is plausible.

One may wonder what use this observation is in a world that is incomprehensibly messy and apparently somewhat non-deterministic. One way of characterizing the insight is to say that *even if* the world were classical, deterministic, very simple, and well-described by highly idealized models—*even if* the world were much more tame than it is—there might *still* be simple questions about it far beyond us to answer, far beyond the reach of any logic or calculation. Hence we should expect complexity and undecidability in our real world all the more, and even in its tidier parts.

To put it another way, even if there is messiness, chaos, and non-determinism in our world, there is also another kind of unpredictability in it that we should understand, obscured though it may be by all the noise. There is *computational* unpredictability, cases where, even given all the relevant information, we just cannot work out the answers we seek. If we want to grasp our world as fully as possible, we need to understand not only noise and non-determinism, but any inherent non-computability that might coexist with them.

However, even if there is no genuine undecidability buried beneath the noise of the real world, let us emphasize again, we have come a long way in clarifying what this might mean—that is, what undecidability in the physical world or in a continuous space could reasonably be, and what it could not be. We have achieved a central goal of philosophy: to explicate the problem.

The reader's patience has been tested to extremes, but perhaps a project of this scale justifies a little indulgence in pomp. Let us conclude by mentioning just a couple of the very broadest morals of our tale.

The approach we have taken here to the questions, "What is a decidable set of real numbers?" and "What is undecidability in the context of a physical system?" has been a pluralistic and pragmatic one. We have argued for the special value of particular concepts, especially d.m.z., but we have always tried to replace "What is" questions with, "What do we want to know or do, and what concepts serve those goals?" We have taken mathematical aesthetics into consideration, in particular a taste for absoluteness or totality, like that of the classical notion of decidability, but not to the exclusion of usefulness.

This pragmatic attitude, and the concept of d.m.z. in particular, can be located in a tradition that Goroff traces to Poincaré:

> There is another approach to the study of dynamical systems which we can also view as having been abstracted from Poincaré's formulation of the recurrence theorem.[5] Less a discipline than a philosophy, it concerns the importance of allowing exceptions. Poincaré's assertion does not concern *all* points in phase space, but rather almost all of them in the sense of measure theory.... The strategy of seeking to describe only typical behavior in these ways has proven particularly fruitful in dynamical systems where, as Poincaré's work illustrates, the set of all possibilities is often too complicated to admit effective classification. (1993, I90)

Poincaré's flexible and permissive approach accords well with what we stated as

[5] A reader unfamiliar with the recurrence theorem should not be concerned here with what it is.

one of the main purposes for pursuing undecidability results: they direct research programs. They tell us what we cannot achieve by particular means and thus on what goals or methods we should not waste sweat. Yet, they do not force us to a dead stop. When it was proved that the circle could not be squared nor the angle trisected with compass and straightedge, mathematicians did not cease to try *altogether*, but only to try by *that method.* They carried on with already established *approximative* methods (not unlike those involved in the concepts of computable real numbers and functions) and to this day continue with an exercise equivalent to squaring the circle, namely computing the digits of π.

Hence, our final moral is this: When we come to an impenetrable rock face let us recognize it as such. Let us not vainly attempt to tunnel through, but nor let us turn back in despair. Let us instead change directions. Perhaps we can in some sense climb upwards, and from a higher vantage point, see things we have not seen before.

APPENDIX

PROOFS CONCERNING THE DISCRETE DYNAMICAL SYSTEM OF OTT ET AL. 1994

It would be easy to lose sight of our central concerns through the formalisms ahead. We primarily wish to prove that the dynamical system we are about to define has two basins of attraction that are not decidable up to measure zero, but are recursively approximable. These basins are sets of *initial conditions*—in this case, *points in the plane*—and the properties that we wish to establish—non-d.m.z. and r.a.—critically involve the *measures* of sets in the plane. In what follows, much is said about *probabilities*, in particular probabilities that an *orbit* will exhibit some kind of behavior. However, probabilities in the mathematical sense are just measures, and the probabilities we will discuss are really just the *Lebesgue measures* of sets of initial conditions or points in the plane. We will speak of them as probabilities only to avail ourselves of some convenient results from probability theory. The risk of confusion is especially great because on a broader motivational level, this dissertation *is* concerned with probabilities that orbits will exhibit some behavior, and the measure-theoretic properties of basins are just tools to understand *those*. However, the *results* we want to prove *in this appendix* are about the Lebesgue measure (and computability) of sets of points, and the probabilities of sets of orbits are just tools to understand that: the *Lebesgue measure* of *sets of points*. It may require some effort on the part of the reader to keep this in mind.

We begin by defining a slight variation on one of the maps from Ott et al. 1994 (407) that is alluded to in Sommerer and Ott's (1996) argument for the validity of their numerical results. Ours is primarily a *generalization* of their map, but we also impose mild computability conditions. Figure 4.5 in Chapter 4 illustrates the definition.

Definition A.1. (i) Let $X = [0, 1] \times [-1, 1]$. Fix computable numbers $\alpha_+ \in (0, 1/2)$, and $\alpha_- \in (1/2, 1)$ and a computable function α: $[-1, 1] \to (0, 1)$ such that for some $y^* > 0$,

$$\alpha(y) = \begin{cases} \alpha_+ \text{ if } y > y^*, \\ \alpha_- \text{ if } y < -y^*. \end{cases}$$

Let f: $[-1, 1] \to [-1, 1]$ be a computable bijection such that for every $y \in (-1, 1)$ we have $f(y) > y$, $\lim_{i \to \infty} f^i(y) = 1$, and $\lim_{i \to \infty} f^{-i}(y) = -1$. Given $(x_0, y_0) \in X$, let $\varphi(x_n, y_n) = (x_{n+1}, y_{n+1})$ for all $n \in \mathbb{N}$, where

$$x_{n+1} = \begin{cases} x_n / \alpha(y_n) & \text{if } x_n < \alpha(y_n), \\ \dfrac{x_n - \alpha(y_n)}{1 - \alpha(y_n)} & \text{otherwise,} \end{cases}$$

and

$$y_{n+1} = \begin{cases} f^{-1}(y_n) & \text{if } x_n < \alpha(y_n), \\ f(y_n) & \text{otherwise.} \end{cases}$$

(ii) The *orbit* of any initial state $(x_0, y_0) \in X$ is a sequence $\xi(x_0, y_0) = \{(x_n, y_n)\} =$

$\{\varphi^n(x_0, y_0)\}_{n \in \mathbb{N}}$.

Notice that any property $P[\xi(x_0, y_0)]$ of an orbit or of its specific elements x_n, y_n defines a set of initial conditions $\{(x_0, y_0): P[\xi(x_0, y_0)]\}$.

We must prove several lemmas in order to obtain our computability and non-computability results. Though Ott et al. supply little detail, the cleverest elements of the proofs (the introduction of probabilities and the reduction of $\{y_n\}_{n \in \mathbb{N}}$ to a random walk) are due to them.

Some of our lemmas involve "coarse-grained histories." The curve $x = \alpha(y)$ partitions the rectangle X into two subsets: $x < \alpha(y)$ and $x \geq \alpha(y)$. A typical orbit will frequently jump from one side of the partition to the other. A *coarse-grained history* is a given pattern of such jumps, which we will represent as a set $H \subseteq \mathbb{N}$ such that $x_n < \alpha(y_n)$ if and only if $n \in H$. Also, for some sets $S \subseteq [0, 1)$ we will be concerned with the set $S_{y_0,H,m}$ consisting of the initial x-coordinates for all orbits $\xi(x_0, y_0)$ that (i) begin at a particular given y_0, (ii) follow a given coarse-grained history H for the first m steps, and (iii) place x_m in a given set $S \subseteq [0, 1)$. More explicitly,

Definition A.2. Let $H \subseteq \mathbb{N}$ and $m \in \mathbb{N}$.

(i) Let $H_m = \{\xi(x_0, y_0): \forall n < m\ [x_n < \alpha(y_n) \Leftrightarrow n \in H]\}$,

(ii) For any $S \subseteq [0, 1)$, and $y_0 \in [-1, 1]$, let

$$S_{y_0,H,m} = \{x_0 \in [0, 1]: \xi(x_0, y_0) \in H_m,\ x_m \in S\}.$$

Our first lemma establishes two facts: that any such set $S_{y_0,H,m}$ is just a linearly squeezed image of S, and that (trivially) all of the orbits with initial conditions in $S_{y_0,H,m} \times \{y_0\}$ have the same y-value after m steps. That is,

Lemma A.3. For each $y_0 \in [-1, 1]$, $H \subseteq \mathbb{N}$, and $m \in \mathbb{N}$, there are constants $a \leq \max_{y \in [-1, 1]}(\max\{\alpha(y)^m, [1 - \alpha(y)]^m\})$, $b < 1$, and $c \in [-1, 1]$ such that for any non-empty set $S \subseteq [0, 1)$,

(i) $S_{y_0,H,m} = aS + b$, and

(ii) $\varphi^m[S_{y_0,H,m} \times \{y_0\}] = S \times \{c\}$.

(Note that a, b, and c are independent of S.)

Proof. We argue by induction on m. The case $m = 0$ is trivial, since $S_{y_0,H,0} = S$. Assume the lemma holds for some particular y_0, H, and m, with constants a, b, and c. This implies that for that particular y_0, if $\xi(x_0, y_0) \in H_m$ then $y_m = c$. We now prove the lemma for y_0, H, and $m + 1$. Let $S \subseteq [0, 1)$.

Case 1: $m \notin H$. Then $\xi(x_0, y_0) \in H_{m+1} \Rightarrow x_m \geq \alpha(y_m)$, so

$$\begin{aligned} S_{y_0,H,m+1} &= \{x_0 \in [0, 1]\colon\ \xi(x_0, y_0) \in H_{m+1}, x_{m+1} \in S\} \\ &= \{x_0 \in [0, 1]\colon\ \xi(x_0, y_0) \in H_m, x_m \geq \alpha(y_m), x_{m+1} \in S\} \\ &= \{x_0 \in [0, 1]\colon\ \xi(x_0, y_0) \in H_m,\ x_m \geq \alpha(y_m),\ \frac{x_m - \alpha(c)}{1-\alpha(c)} \in S\} \end{aligned}$$

$$= \{x_0 \in [0, 1]:\ \xi(x_0, y_0) \in H_m,\ x_m \in [1 - \alpha(c)]S + \alpha(c)\}$$

$$= a(\,[1 - \alpha(c)]S + \alpha(c)\,) + b \quad \text{(by induction hypothesis)}$$

$$= a[1 - \alpha(c)]S + [a\alpha(c) + b]$$

$$= a'S + b',$$

where $a' = a[1 - \alpha(c)]$ and $b' = [a\alpha(c) + b]$. Also,

$$a' \leq \max_{y \in [-1, 1]}(\max\{\alpha(y)^m, [1 - \alpha(y)]^m\}) \cdot \max_{y \in [-1, 1]}[1 - \alpha(y)]$$

$$< \max_{y \in [-1, 1]}(\max\{\alpha(y)^m, [1 - \alpha(y)]^m\}),$$

and $b' < 1$, since $S_{y_0,H,m+1} \subseteq [0, 1] \times \{y_0\}$. Lastly,

$$\varphi^{m+1}[S_{y_0,H,m+1} \times \{y_0\}] = \varphi\varphi^m[\{x_0\!:\ \xi(x_0, y_0) \in H_m,\ x_{m+1} \in S\} \times \{y_0\}]$$

$$= \varphi\varphi^m[\{x_0\!:\ \xi(x_0, y_0) \in H_m,\ x_m \in [1 - \alpha(c)]S + \alpha(c)\} \times \{y_0\}]$$

$$= \varphi[(\,[1 - \alpha(c)]S + \alpha(c)\,) \times \{c\}] \qquad \text{(by hypothesis)}$$

$$= S \times \{c'\},$$

where $c' = f(c)$.

Case 2: $m \in H$. The argument is parallel but a little simpler, with $a' = a\alpha(c)$, $b' = b$, and $c' = f^{-1}(c)$. ■

We now introduce a probability. Again, we are mainly concerned here with subsets of the rectangle X and their two-dimensional Lebesgue measure. However (and this was not said above), we will first have to prove some things about the *one-*

dimensional Lebesgue measure of certain subsets of *horizontal line segments*. It is this one-dimensional Lebesgue measure on each horizontal that we treat as a probabiltiy. Thus we actually define a *family* $\{\Pr_{y_0}\}_{y_0 \in [-1,1]}$ of probability measures $\Pr_{y_0}$, each defined on sets of orbits originating at $y = y_0$. An expression of the from $\Pr_c[S]$ may be thought of as representing $\Pr[S \mid y_0 = c]$, where 'Pr' denotes a probability measure on *unrestricted* sets of orbits that do *not* necessarily share a common y_0, equivalent to the *two*-dimensional Lebesgue measure on the corresponding sets of initial conditions. However, we do not need to define the more general probability Pr, for we make no use of it. Instead, we first convert our *restricted* probabilistic results back into measure-theoretic results for subsets of line segments, and then integrate these measure-theoretic results to obtain the two-dimensional measures in $\mathbb{R}^2$ that we ultimately want.

Our family of probability measures is defined as follows:

Definition A.4. (i) Let λ_1 denote the Lebesgue measure on $\mathbb{R}$.

(ii) Let $y_0 \in [-1, 1]$, and let Γ and Δ be sets of orbits such that $\{x_0 \in [0, 1]\colon \xi(x_0, y_0) \in \Gamma\}$ and $\{x_0 \in [0, 1]\colon \xi(x_0, y_0) \in \Delta\}$ are Lebesgue measurable and the latter has non-zero measure. Then

$$\Pr_{y_0}[\Gamma] = \lambda_1\{x_0 \in [0, 1]\colon \xi(x_0, y_0) \in \Gamma\}, \quad \text{and}$$

$$\Pr_{y_0}[\Gamma \mid \Delta] = \frac{\lambda_1\{x_0 \in [0,1]\colon \xi(x_0, y_0) \in \Gamma \cap \Delta\}}{\lambda_1\{x_0 \in [0,1]\colon \xi(x_0, y_0) \in \Delta\}}.$$

Formally, these probabilities are defined on sets of orbits. However, we often express a probability in terms of some property of the elements of orbits, such as '$x_m \in S$' or '$\exists n\ |y_n| \geq c$,' which of course defines a set of orbits. This set of orbits in turn defines a set of initial x-values, and the probability of the condition is equal to the Lebesgue measure of the latter set. For the conditions we will consider here, these sets of initial x-values are all measurable, as the reader may verify in two ways: we define them without appeal to the Axiom of Choice or any equivalent, and they are all Borel (i.e., they can be generated by countably many operations of union and intersection on open and closed sets).

The next lemma, a direct corollary of the preceding one, shows that in a certain sense the future of an orbit is independent of the orbit's coarse-grained history.

Lemma A.5 (Independence Lemma). Let $y_0 \in [-1, 1]$, $m \in \mathbb{N}$, $H \subseteq \mathbb{N}$, and S a measurable subset of $[0, 1)$. Then regardless of H, $\Pr_{y_0}[x_m \in S \mid \xi(x_0, y_0) \in H_m] = \lambda_1 S$.

Proof. By Lemma A.3, there exist constants a and b such that

$$\Pr_{y_0}[x_m \in S \mid \xi(x_0, y_0) \in H_m] = \frac{\lambda_1\{x_0 \in [0,1] : x^m \in S, \xi(x_0, y_0) \in H_m\}}{\lambda_1\{x \in [0,1] : \xi(x_0, y_0) \in H_m\}}$$

$$= \frac{\lambda_1 S_{y_0,H,m}}{\lambda_1(\,[0,1)_{y_0,H,m})}$$

$$= \frac{\lambda_1[aS + b]}{\lambda_1[a[0,1) + b]}$$

$$= \frac{\lambda_1 S}{1} = \lambda_1 S . \blacksquare$$

The next lemma says that the probability that the sequence $\{y_n\}$ will depart from a given bounded region within an allotted time converges effectively to 1 as the allotted time is increased.

Lemma A.6. Let $y_0 \in [-1, 1]$, $c \in [0, 1)$, and $N, M, k \in \mathbb{N}$ such that $N < M$ and $f^k(0)$, $|f^{-k}(0)| > c$. Then

$$\Pr_{y_0}[(\exists n \in \mathbb{N} \cap (N, M])\ |y_n| \geq c] > 1-(1-[\min_{y\in[-1,1]} \alpha(y)]^{2k})^{\frac{M-N}{2k}-1}.$$

Proof. If for some $m \in \mathbb{N} \cap (N, M-2k]$, $x_m < \alpha(y_m)$, $x_{m+1} < \alpha(y_{m+1}), \ldots, x_{m+2k-1} < \alpha(y_{m+2k-1})$, then either $y_m > f^k(0)$ or $y_{m+2k} < f^{-k}(0)$. In either case, we have $n \in \mathbb{N} \cap (N, M]$ such that $|y_n| \geq c$. Therefore it is enough to show that (letting n, m range over $\mathbb{N}$)

$$\Pr_{y_0}[(\exists m \in (N, M-2k])(\forall n \in [m, m+2k))\ x_n < \alpha(y_n)]$$

$$> 1-(1-[\min_{y\in[-1,1]} \alpha(y)]^{2k})^{\frac{M-N}{2k}-1}.$$

This probability is analogous to that of rolling a 1 on a die $2k$ consecutive times within $M - N$ independent trials. To obtain a coarse lower bound, we may divide the $M-N$ trials into $\lfloor (M-N)/2k \rfloor$ sequences (where $\lfloor \cdot \rfloor$ denotes the greatest-integer function), each consisting of $2k$ trials, and ignore any remainder:

$$\Pr_{y_0}[(\exists m \in (N, M-2k])(\forall n \in [m, m+2k))\ x_n < \alpha(y_n)]$$

$$\geq \Pr_{y_0}\left[(\exists m < \left\lfloor \frac{M-N}{2k} \right\rfloor)(\forall n \in (N+2km, N+2km+2k])\ x_n < \alpha(y_n)\right]$$

$$= 1 - \Pr_{y_0}\left[(\forall m < \left\lfloor \frac{M-N}{2k} \right\rfloor)(\exists n \in (N+2km, N+2km+2k])\ x_n \geq \alpha(y_n)\right]$$

$$= 1 - \prod_{m=0}^{\left\lfloor \frac{M-N}{2k} \right\rfloor - 1} \Pr_{y_0}\left[(\exists n \in (N+2km, N+2km+2k])\, x_n < \alpha(y_n) \,\middle|\, (\forall i < m)(\exists n \in [2ki, 2ki+2k))\, x_n \geq \alpha(y_n)\right]$$

$$= 1 - \prod_{m=0}^{\left\lfloor \frac{M-N}{2k} \right\rfloor - 1} \Big(1 - \prod_{n=N+2km+1}^{N+2km+2k} \Pr_{y_0}\left[x_n < \alpha(y_n) \,\middle|\, (\forall j \in [2km, n))\, x_j > \alpha(y_j) \ \&\ (\forall i < m)(\exists n \in [2ki, 2ki+2k))\, x_n \geq \alpha(y_n)\right]\Big)$$

$$\geq 1 - \prod_{m=0}^{\left\lfloor \frac{M-N}{2k} \right\rfloor - 1} \Big(1 - \prod_{n=N+2km+1}^{N+2km+2k} \Pr_{y_0}\left[x_n < \min_{y \in [-1,1]} \alpha(y) \,\middle|\, (\forall j \in [2km, < n))\, x_j > \alpha(y_j) \ \&\ (\forall i < m)(\exists n \in [2ki, 2ki+2k))\, x_n \geq \alpha(y_n)\right]\Big)$$

$$= 1 - \prod_{m=0}^{\left\lfloor \frac{M-N}{2k} \right\rfloor - 1} \Big(1 - \prod_{n=N+2km+1}^{N+2km+2k} \lambda_1[0, \min_{y \in [-1,1]} \alpha(y))\Big) \quad \text{(by Independence Lemma)}$$

$$= 1 - \left(1 - [\min_{y \in [-1,1]} \alpha(y)]^{2k}\right)^{\left\lfloor \frac{M-N}{2k} \right\rfloor}$$

$$> 1-(1-[\min_{y\in[-1,1]} \alpha(y)]^{2k})^{\frac{M-N}{2k}-1}. \blacksquare$$

It follows quickly that almost every orbit will spend an infinite amount of time outside any region $[0, 1] \times (-c, c)$ where $c < 1$. That is,

Lemma A.7. Let $y_0 \in [-1, 1]$, $c \in [0, 1)$. Then $\Pr_{y_0}[(\forall N)(\exists n > N)|y_n| > c] = 1$.

Proof. $\Pr_{y_0}[(\forall N)(\exists n > N)\,|y_n| \geq c]$

$$= \prod_{N=1}^{\infty} \Pr_{y_0}\left[(\exists n > N)\,|y_n| \geq c \mid (\forall M < N)(\exists n > N)\,|y_n| > c\right]$$

$$\geq \prod_{N=1}^{\infty} \Pr_{y_0}[(\exists n > N)\,|y_n| \geq c]$$

$$\geq \prod_{N=1}^{\infty} \lim_{M\to\infty} \Pr_{y_0}[(\exists n \ni N < n \leq M)\,|y_n| \geq c]$$

$$= \prod_{N=1}^{\infty} 1 = 1 \qquad \text{(by the preceding lemma).} \blacksquare$$

At this point we make use of the notion of a simple random walk.

Definition A.8. A *simple random walk* is a finite or infinite sequence $\{z_m\}$ of random variables z_m such that for each m, $z_m = z_0 + Z_1 + \ldots + Z_m$, where the Z_i are independent and identically distributed random variables taking values in $\{-1, 0, 1\}$.

Intuitively, coordinates y_n greater than y^* or less than $-y^*$ behave as a simple random walk over the values $\{f^i(y_0)\}_{i \in \mathbb{Z}}$. But according to the usual definition, the one just given, a random walk jumps only by integers. To meet this definition on its own terms, we define a sequence $\{z_m\}$ of integers that rises by one unit if $y_n = f(y_{n-1})$ and falls by one unit if $y_n = f^{-1}(y_{n-1})$. The precise definition of $\{z_m\}$ is given in the following lemma.

Lemma A.9. Given $y_0 \in [-1, 1]$, and $n, N \in \mathbb{N}$ such that $n > N$, let $z_0 = 0$, and for each $m \in \{1, 2,\ldots, n-N\}$ define Z_m by $y_{N+m} = f^{Z_m}(y_{N+m-1})$, and let $z_m = z_0 + Z_1 + \ldots + Z_m$. Then...

(i) If $y_N, y_{N+1},\ldots, y_n \geq y^*$ then $z_0, z_1,\ldots, z_{n-N}$ is a simple random walk with $\Pr_{y_0}[Z_m = -1] = \alpha_+$ and $\Pr_{y_0}[Z_m = 1] = 1 - \alpha_+$.

(ii) If $y_N, y_{N+1},\ldots, y_n \leq -y^*$ then $z_0, z_1,\ldots, z_{n-N}$ is a simple random walk with $\Pr_{y_0}[Z_m = -1] = \alpha_-$ and $\Pr_{y_0}[Z_m = 1] = 1 - \alpha_-$.

Proof. Suppose $y_N, y_{N+1},\ldots, y_n \geq y^*$. We need to show that the Z_m are independent and all have the stated distribution. Notice that for any $m \in \{1, 2,\ldots, N-n\}$ and $i_1, i_2,\ldots, i_{m-1} \in \{-1, 1\}$,

$$\Pr_{y_0}[Z_m = -1 \mid Z_1 = i_1, Z_2 = i_2, \ldots, Z_{m-1} = i_{m-1}]$$

$$= \Pr_{y_0}[y_{N+m} = f^{-1}(y_{N+m-1}) \mid \forall j \in \{1,\ldots,m-1\}(y_{N+j} = f^{i_j}(y_{N+j-1}))]$$

$$= \Pr_{y_0}[x_{N+m-1} < \alpha_+ \mid \forall j \in \{1,\ldots,m-1\}(x_{N+j-1} < \alpha_+ \Leftrightarrow i_j = -1)].$$

Let $\mathbf{H} = \{H \subseteq \mathbb{N} \mid \forall j \in \{1,\ldots, m-1\}(N+j-1 \in H \Leftrightarrow i_j = -1)\}$. Then by the above,

$$\Pr_{y_0}[Z_m = -1 \mid Z_1 = i_1, Z_2 = i_2, \ldots Z_{m-1} = i_{m-1}]$$

$$= \Pr_{y_0}[x_{N+m-1} < \alpha_+ \mid (\exists H \in \mathbf{H})\xi(x_0, y_0) \in H_{N+m-2}].$$

It follows trivially from the Independence Lemma that this is equal to $\Pr_{y_0}[x_{N+m-1} < \alpha_+] = \alpha_+$. But also, $\Pr_{y_0}[x_{N+m-1} < \alpha_+] = \Pr_{y_0}[Z_m = -1]$, so

$$\Pr_{y_0}[Z_m = -1 \mid Z_1 = i_1, Z_2 = i_2, \ldots Z_{m-1} = i_{m-1}] = \Pr_{y_0}[Z_m = -1] = \alpha_+.$$

Similarly,

$$\Pr_{y_0}[Z_m = 1 \mid Z_1 = i_1, Z_2 = i_2, \ldots Z_{m-1} = i_{m-1}] = \Pr_{y_0}[Z_m = 1] = 1 - \alpha_+.$$

The case $y_N, y_{N+1}, \ldots, y_n \leq -y^*$ is parallel. ■

The next lemma is a well-known fact about simple random walks. It implies that the probability that a walk will ever return to its starting point decreases to zero in a very regular (and computable) way as the walk wanders farther away.

Lemma A.10. Let $\{z_m\}$ be a simple random walk, with $\Pr_{y_0}[Z_m = -1] = \gamma$ and $\Pr_{y_0}[Z_m = 1] = 1 - \gamma$. Let $C = z_0 + i$ for some $i \in \mathbb{Z}^+$. Then

$$\Pr_{y_0}[(\forall m \geq M)\, z_m \geq C \mid z_M = C + k] = 1 - \left(\frac{\gamma}{1-\gamma}\right)^k,$$

$$\Pr_{y_0}[(\forall m \geq M)\, z_m \leq C \mid z_M = C - k] = 1 - \left(\frac{1-\gamma}{\gamma}\right)^k.$$

Proof. See for example Hoel, Port, and Stone 1971, 222-223.

The following corollary is, among other things, the key to our proof that the basins are recursively approximable (Proposition A.21). It implies that the probability that an orbit beginning at $y_0 = y$ tends in the limit to a given attractor decreases to zero in a very regular way as the orbit comes close to the *other* attractor. Here $f^{k-1}(c) \leq y_N < f^k(c)$ expresses the condition that y_N is k "steps" (iterations of f) above c, and $f^{-k}(-c) < y_N \leq f^{-k+1}(-c)$ says that y_N is k steps below $-c$. This is followed by a simpler but also useful corollary.

Corollary A.11. Let $y_0 \in [-1, 1]$, $c \in [y^*, 1)$, $N \in \mathbb{N}$. Then

$$\Pr_{y_0}[(\forall n \geq N)\, y_n \geq c \mid f^{k-1}(c) \leq y_N < f^k(c)] = 1 - \left(\frac{\alpha_+}{1-\alpha_+}\right)^k, \text{ and}$$

$$\Pr_{y_0}[(\forall n \geq N)\, y_n \leq -c \mid f^{-k}(-c) < y_N \leq f^{-k+1}(-c)] = 1 - \left(\frac{1-\alpha_-}{\alpha_-}\right)^k.$$

Proof. Immediate from Lemmas A.9 and A.10.

Corollary A.12. Let $y_0 \in [-1, 1]$, $c \geq y^*$, $N \in \mathbb{N}$. Then

$$\Pr_{y_0}[(\forall n \geq N)\, |y_n| \geq c \mid |y_N| \geq c] \geq \min\left\{\frac{1-2\alpha_+}{1-\alpha_+}, \frac{2\alpha_- - 1}{\alpha_-}\right\}.$$

Proof. By the previous corollary,

$$\Pr_{y_0}[(\forall n \geq N)\, y_n \geq c \mid y_N \geq c] \geq 1 - \left(\frac{\alpha_+}{1-\alpha_+}\right) = \frac{1-2\alpha_+}{1-\alpha_+}, \text{ and}$$

$$\Pr_{y_0}[(\forall n \geq N)\, y_n \leq -c \mid y_N \leq -c] \geq 1 - \left(\frac{1-\alpha_-}{\alpha_-}\right) = \frac{2\alpha_- - 1}{\alpha_-}.$$

Therefore,

$$\Pr_{y_0}[(\forall n \geq N)\, |y_n| \geq c \mid |y_N| \geq c]$$

$$= \Pr_{y_0}[(\forall n \geq N)\, y_n \geq c \mid |y_N| \geq c] + \Pr_{y_0}[(\forall n \geq N)\, y_n \leq -c \mid |y_N| \geq c]$$

$$= \Pr_{y_0}[(\forall n \geq N)\, y_n \geq c \mid y_N \geq c]\ \Pr_{y_0}[y_N \geq c \mid |y_N| \geq c]$$

$$+ \Pr_{y_0}[(\forall n \geq N)\, y_n \leq -c \mid y_N \leq -c]\ \Pr_{y_0}[y_N \leq -c \mid |y_N| \geq c]$$

$$\geq \min\left\{\frac{1-2\alpha_+}{1-\alpha_+}, \frac{2\alpha_- - 1}{\alpha_-}\right\} \Pr_{y_0}[y_N \geq c \mid |y_N| \geq c]$$

$$+ \min\left\{\frac{1-2\alpha_+}{1-\alpha_+}, \frac{2\alpha_- - 1}{\alpha_-}\right\} \Pr_{y_0}[y_N \leq -c \mid |y_N| \geq c]$$

$$= \min\left\{\frac{1-2\alpha_+}{1-\alpha_+}, \frac{2\alpha_- - 1}{\alpha_-}\right\}. \blacksquare$$

The next lemma says that almost every orbit in a given horizontal is attracted to one or the other extreme horizontal, $A_+ = [0, 1] \times \{1\}$ or $A_- = [0, 1] \times \{-1\}$.

Lemma A.13. $\Pr_{y_0}[\{y_n\} \to 1 \text{ or } \{y_n\} \to -1] = 1$.

Proof. The idea is this: We know from Lemma A.7 that for any $c < 1$, it is virtually certain that for infinitely many n, $|y_n| \geq c$. For each such n there is a significant chance that for $m \geq n$, $|y_m|$ will never again fall below c (Corollary A.12). Therefore, intuitively, the sequence $\{y_n\}$ should eventually escape any interval $(-c, c)$ forever.

To be precise, fix an infinite sequence $\{c_i\}_{i \in \mathbb{N}}$ of numbers in $(y^*, 1)$ such than $\{c_i\} \to 1$. Then

$$\Pr_{y_0}[\{y_n\} \to 1 \text{ or } \{y_n\} \to -1] = \Pr_{y_0}[(\forall i)(\exists N)(\forall n \geq N)\, |y_n| \geq c_i]$$

$$= \prod_{i=0}^{\infty} \Pr_{y_0}[(\exists N)(\forall n \geq N)\, |y_n| \geq c_i \mid (\forall j < i)(\exists N)(\forall n \geq N)\, |y_n| \geq c_j]$$

$$\geq \prod_{i=0}^{\infty} \Pr_{y_0}[(\exists N)(\forall n \geq N)\, |y_n| \geq c_i].$$

Therefore it is enough to show that for each c_i, $\Pr_{y_0}[(\exists N)(\forall n \geq N)\, |y_n| \geq c_i] = 1$.

To see that, fix $c_i \geq y^*$. Note $\Pr_{y_0}[(\exists N)(\forall n \geq N)\, |y_n| \geq c_i] = 1 - \Pr_{y_0}$ [for infinitely many n, $|y_n| < c_i$]. By Lemma A.7, $\Pr_{y_0}$ [for infinitely many n, $|y_n| \geq c_i$] = 1, so $\Pr_{y_0}$ [for infinitely many n, $|y_n| < c_i$] = $\Pr_{y_0}$ [$\{y_n\}$ escapes $(-c_i, c_i)$ and returns again infinitely many times]. We want to show that this probability is 0.

Consider $\Pr_{y_0}$ [$\{y_n\}$ escapes $(-c_i, c_i)$ and returns again at least N times]. This equals

$$\prod_{j=1}^{N} \Big(\Pr_{y_0} [\{y_n\} \text{ escapes } (-c_i, c_i) \text{ and returns again at least } j \text{ times} \mid \{y_n\} \text{ escapes } (-c_i, c_i) \text{ at least } j \text{ times}]$$

$$\cdot \Pr_{y_0} [\{y_n\} \text{ escapes } (-c_i, c_i) \text{ at least } j \text{ times} \mid \{y_n\} \text{ escapes } (-c_i, c_i) \text{ and returns at least } j-1 \text{ times}] \Big).$$

By Lemmas A.5 (Independence) and A.12, the first probability in this product is less than or equal to $\max\{\alpha_+ / (1 - \alpha_+), (1 - \alpha_-) / \alpha_-\}$, and of course the second is less than or equal to 1. Therefore,

$$\Pr_{y_0} [\{y_n\} \text{ escapes } (-c_i, c_i) \text{ and returns again at least } N \text{ times}]$$

$$\leq \prod_{i=1}^{N} \max\left\{\frac{\alpha_+}{1-\alpha_+}, \frac{1-\alpha_-}{\alpha_-}\right\} = \max\left\{\frac{\alpha_+}{1-\alpha_+}, \frac{1-\alpha_-}{\alpha_-}\right\}^N,$$

so

$$\Pr_{y_0} [\{y_n\} \text{ escapes } (-c_i, c_i) \text{ and returns again infinitely many times}]$$

$$\leq \lim_{N\to\infty} \max\left\{\frac{\alpha_+}{1-\alpha_+}, \frac{1-\alpha_-}{\alpha_-}\right\}^N = 0,$$

since $\alpha_+ < \frac{1}{2} < \alpha_-$. To sum up,

$$\Pr_{y_0} [\{y_n\} \to 1 \text{ or } \{y_n\} \to -1] \geq \prod_{i=0}^{\infty} \Pr_{y_0} [(\exists N)(\forall n \geq N)\, |y_n| \geq c_i]$$

$$= \prod_{i=0}^{\infty} 1 - \Pr_{y_0} [\text{for infinitely many } n, |y_n| < c_i]$$

$$= \prod_{i=0}^{\infty} 1 - \Pr_{y_0} [\{y_n\} \text{ escapes } (-c_i, c_i) \text{ and returns again infinitely many times}]$$

$$= \prod_{i=0}^{\infty} 1 - 0 = 1. \blacksquare$$

This means that for initial conditions on any one horizontal line, the probability that either $\{y_n\} \to 1$ or $\{y_n\} \to -1$ is 1. It follows quickly that almost every initial condition in X lies in the basin of attraction of either A_+ or A_-. This is spelled out in statements A.14-A.17.

Definition A.14. (i) Let λ_2 denote the Lebesgue measure on $\mathbb{R}^2$.

(ii) Let $\beta(A)$ denote the *basin* of A, i.e., the set $\{(x_0, y_0): d[(x_n, y_n), A] \to 0\}$.

Corollary A.15. For any $y_0 \in [-1, 1]$,

$$\lambda_1\{x_0 \in [0, 1]: (x_0, y_0) \in \beta(A_+) \text{ or } (x_0, y_0) \in \beta(A_-)\} = 1.$$

Proof. Immediate from A.14. ■

Lemma A.16. If for each y_0 in an interval $[a, b]$, S_{y_0} is a Lebesgue measurable subset of $\mathbb{R}$ with finite measure and $\{(x_0, y_0) \in \mathbb{R} \times [a, b] : x_0 \in S_{y_0}\}$ is measurable, then

$$\lambda_2\{(x_0, y_0) \in \mathbb{R}\times [a, b] : x_0 \in S_{y_0}\} = \int_a^b \lambda_1 S_{y_0} \mathrm{d}y_0 .$$

Proof. This is a special case of Fubini's Theorem. See, e.g., Folland 1984, 64-65. ■

Proposition A.17. $\lambda_2[\beta(A_+) \cup \beta(A_-)] = 2$ and $\lambda_2(X / [\beta(A_+) \cup \beta(A_-)]) = 0$.

Proof. By A.15 and A.16,

$$\lambda_2[\beta(A_+) \cup \beta(A_-)] = \int_{-1}^{1} \lambda_1\{x_0 : (x_0, y_0) \in \beta(A_+) \cup \beta(A_-)\} \mathrm{d}y_0$$

$$= \int_{-1}^{1} 1 \mathrm{d}y_0 = 2 . \blacksquare$$

We now come to one of our desired results.

Proposition A.18. The basins $\beta(A_+)$ and $\beta(A_-)$ are intermingled.

Proof. We shall see that in any small neighborhood, both basins have positive measure. First we wish to see that any set of the form $[0, 1) \times \{c\}$ contains positive-measure portions of both basins. Choose c. Since almost all orbits on any horizontal go to one attractor or the other, we have for each $N \in \mathbb{N}$,

$$\Pr_c [\{y_n\} \to 1] \geq \Pr_c [\{y_n\} \to 1 \mid (\forall n \geq N)\, y_n \geq y^*] \cdot \Pr_c [(\forall n \geq N)\, y_n \geq y^*]$$

$$= \Pr_c [(\forall n \geq N)\, y_n \geq y^*]$$

$$= \Pr_c[(\forall n \geq N)\, y_n \geq y^* \mid y_N \geq y^*] \cdot \Pr_c[y_N \geq y^*].$$

Clearly $\Pr_c[y_N \geq y^*] > 0$ for some N, and by the proof of Corollary A.12, $\Pr_c[(\forall n \geq N)\, y_n \geq y^* \mid y_N \geq y^*] > 0$. Therefore $\Pr_c[\{y_n\} \to 1] > 0$, and by a parallel argument, $\Pr_c[\{y_n\} \to -1] > 0$. Hence, by definition of $\Pr_c$, $\lambda_1[\beta(A_+) \cap ([0, 1) \times \{c\})] > 0$ and $\lambda_1[\beta(A_-) \cap ([0, 1) \times \{c\})] > 0$.

Next we show that the images, by iterations of φ, of any small horizontal line segment contain some set of the form $[0, 1) \times \{c\}$. That is, for any $y_0 \in [-1, 1]$ and any interval $I \subseteq [0, 1]$ there exists $m \in \mathbb{N}$, $c \in [-1, 1]$, and an interval $J \subseteq I$ such that $\varphi^m[J \times \{y_0\}] = [0, 1) \times \{c\}$. To see this, fix I and choose m such that $\max_{y \in [-1, 1]} \max\{\alpha(y)^m, [1 - \alpha(y)]^m\} < (\lambda_1 I)/2$. Let $S = [0, 1)$. Then by Lemma A.3, for each $H \subseteq \mathbb{N}$ there exist $a < (\lambda_1 I)/2$, $b < 1$, and $c \in [-1, 1]$ such that $S_{y_0,H,m} = (a[0, 1) + b)$ and $\varphi^m[S_{y_0,H,m} \times \{y_0\}] = [0, 1) \times \{c\}$. For fixed y_0 and m but varying $H \subseteq \{0, 1,\ldots, m-1\}$, the sets $S_{y_0,H,m}$ are pairwise disjoint line segments with length less than $(\lambda_1 I)/2$, and every initial condition $(x_0, y_0) \in [0, 1) \times \{y_0\}$ lies in some such $S_{y_0,H,m} \times \{y_0\}$. Therefore $I \times \{y_0\}$ contains at least one segment $S_{y_0,H,m}$ with $\varphi^m[S_{y_0,H,m} \times \{y_0\}] = [0, 1) \times \{c\}$.

Combining these results, we show that each basin has positive measure in every small horizontal line segment. Choose I and y_0 and fix H, c, and m such that $\varphi^m[S_{y_0,H,m} \times \{y_0\}] = [0, 1) \times \{c\}$, where S is still $[0, 1)$. Let $T = \beta(A_+) \cap ([0, 1) \times \{c\})$. By the paragraph before last, $\lambda_1 T > 0$. By Lemma A.3, $T_{y_0,H,m} = [aT + b] \subseteq S_{y_0,H,m} \subseteq$

I. Also, $T_{y_0,H,m} \times \{y_0\} \subseteq \beta(A_+)$, so $T_{y_0,H,m} \times \{y_0\} \subseteq \beta(A_+) \cap (I \times \{y_0\})$. Since $\lambda_1 T_{y_0,H,m} = \lambda_1[aT+b] = a\lambda_1 T > 0$, we have $\lambda_1[\beta(A_+) \cap (I \times \{y_0\})] > 0$, and by a parallel argument, $\lambda_1[\beta(A_-) \cap (I \times \{y_0\})] > 0$.

To complete the proof, let $U \subseteq X$ be open. Then U contains an open disk V, so by Lemma A.16,

$$\lambda_2(U \cap \beta(A_+)) \geq \lambda_2(V \cap \beta(A_+)) = \int_{-1}^{1} \lambda_1[V \cap ([0,1] \times \{y_0\}) \cap \beta(A_+)]\,\mathrm{d}y_0\,,$$

and by the preceding paragraph, this integral is greater than zero. By a parallel argument, $\lambda_2(U \cap \beta(A_-)) > 0$. Therefore $\beta(A_+)$ and $\beta(A_-)$ are intermingled. ■

It follows immediately, by Theorem 3.6.2 of Chapter 3, that the basins are not decidable up to measure zero. Even worse, they are *nowhere* d.m.z. (Definition 3.5.12(iii)).

Proposition A.19. The basins $\beta(A_+)$ and $\beta(A_-)$ are nowhere λ_2-d.m.z.

Proof. Since $\beta(A_+)$ and $\beta(A_-)$ are dense in X and are intermingled, both $U \cap \beta(A_+)$ and $U \cap \beta(A_-)$ are riddled and have positive measure for any open set $U \subseteq \mathbb{R}^2$. Therefore no such set $U \cap \beta(A_+)$ or $U \cap \beta(A_-)$ is d.m.z. ■

Trivially, the basins are not r.e. open either, nor are they r.e. closed, Π^0_1, TTE-recursive, nor strongly recursive, simply because...

Remark A.20. The basins $\beta(A_+)$ and $\beta(A_-)$ are neither open nor closed.

Proof. Every open set contains elements of $\beta(A_+)^C$ and $\beta(A_-)^C$, so the basins are not open. Every open set also contains elements of $\beta(A_+)$ and $\beta(A_-)$, so they are not complements of open sets. ■

Also trivially, the basins *are* Δ-decidable and d.i.b.: Since every open set contains portions of both basins, the distance to the boundary of either basin is always zero. An algorithm that blindly outputs zero therefore computes the delta function for both basins, and also correctly decides the basins except on the boundary—but the boundary is everywhere! This is a case where d.i.b. and Δ-decidability hold, but only due to the intermingling that makes membership thoroughly *un*decidable.

However, all computation is not hopeless, as our final result shows.

Proposition A.21. The basins $\beta(A_-)$ and $\beta(A_+)$ are recursively approximable.

Proof. Our strategy is to approximate the orbit of a given point until it comes very close to one of the attractors, much as Sommerer and Ott did for their continuous-time system (1996). We know from Corollary A.11 that if an orbit does come very close to an attractor, there is very little chance that it eventually tends toward the other attractor. Since almost all orbits tend toward one attractor or the other, that means there is a great chance that the orbit lies in the basin of the nearby attractor.

We must still handle some details, such as ensuring that our algorithm always halts. The algorithm involves comparing real numbers, and exact comparisons cannot always be made in finite time. Therefore we dovetail each comparison with another in such a way that one or the other will halt.

Let a be a computable number no greater than $\min_{y \in [-1,1]} \alpha(y)$, and let $y^{*\prime}$ be a computable number greater than y^*. An algorithm to approximate $\beta(A_+)$ given input (ϕ, i), where $\rho_{\mathbb{R}^2}(\phi) = (x_0, y_0)$, is as follows:

1. Let $n = 0$.
2. Dovetail the following two comparisons: If $|x_n - \alpha(y_n)| < 2^{-2n-i-5}$ then output 0 and halt; if $|x_n - \alpha(y_n)| > 2^{-2n-i-6}$, proceed to step 3.
3. Find an integer $j > \log 2^{-i-3} / \log[\alpha_+ / (1 - \alpha_+)]$, and dovetail again: If $y_n > f^j(y^{*\prime})$, output 1 and halt; if $y_n < f^{j+1}(y^{*\prime})$, proceed to step 4.
4. Find an integer $k > \log 2^{-i-3} / \log[(1 - \alpha_-) / \alpha_-]$ and dovetail: If $y_n < f^{-k}(-y^{*\prime})$, output 0 and halt; if $y_n > f^{-k-1}(-y^{*\prime})$, proceed to step 5.
5. Find an integer m such that $f^{-m}(0) < f^{-k}(-y^{*\prime})$ and $f^m(0) > f^j(y^{*\prime})$, and dovetail: If $n > [2m \log 2^{-i-3} / \log(1 - a^{2m})] + 2m$, output 0 and halt; if $n < [2m \log 2^{-i-3} / \log(1 - a^{2m})] + 2m + 1/2$, set $n = n + 1$ and go to step 2.

Step 2 circumvents any difficulty in approximating (x_n, y_n) arising from the discontinuity of φ along the curve $x = \alpha(y)$: if x_n is near $\alpha(y_n)$, the algorithm halts. This can result in incorrect output, but on any given horizontal line $y = y_0$, the errors due to step 2 are confined to small neighborhoods near the pre-images by φ^n of the curve $x =$

$\alpha(y)$. For each n, the pre-image by φ^n of $x = \alpha(y)$ intersects a horizontal in 2^n points, so the total measure of the neighborhoods around these points where step 2 may yield incorrect output is less than

$$\sum_{n=0}^{\infty} 2^n 2 \cdot 2^{-2n-i-5} = 2^{-i-3}.$$

Now consider the other steps. By Corollary A.11, the measure of the set of all points on $y = y_0$ that are incorrectly classified by step 3 is less than

$$[\alpha_+ /(1-\alpha_+)]^{\log 2^{-i-3} / \log[\alpha_+ /(1-\alpha_+)]} = 2^{-i-3}.$$

Similarly, the set of points incorrectly classified by step 4 has measure less than $[(1-\alpha_-)/\alpha_-]^{\log 2^{-i-3} / \log[(1-\alpha_-)/\alpha_-]} = 2^{-i-3}$. By Lemma A.6, the set of points on $y = y_0$ that are classified (perhaps incorrectly) by step 5, i.e., the set of points whose orbits do not exit the region $[0, 1] \times [f^{-k-1}(-y^{*\prime}), f^{j+1}(y^{*\prime})]$ within $\lfloor [2m \log 2^{-i-3} / \log(1 - a^{2m})] + 2m + 1 \rfloor$ iterations of φ, has measure less than

$$\begin{aligned}
\left(1-a^{2m}\right)^{\frac{\left\lfloor \frac{2m \log 2^{-i-3}}{\log(1-a^{2m})} + 2m+1 \right\rfloor}{2m} - 1} &\le \left(1-a^{2m}\right)^{\frac{\frac{2m \log 2^{-i-3}}{\log(1-a^{2m})} + 2m}{2m} - 1} \\
&= \left(1-a^{2m}\right)^{\log 2^{-i-3} / \log(1-a^{2m})} \\
&= 2^{-i-3}.
\end{aligned}$$

Hence the total measure of the set of *all* points on $y = y_0$ that are incorrectly classified by one step or another, given the parameter i, is less than $4(2^{-i-3}) = 2^{-i-1}$. By

Lemma A.16, the total *two-dimensional* Lebesgue measure of the set of all points in X incorrectly classified by this algorithm is therefore less than $2(2^{-i-1}) = 2^{-i}$. The algorithm will halt on every input because step 5 sets a finite bound on n (the number of repetitions). Therefore $\beta(A_+)$ is recursively approximable, and a parallel argument shows that $\beta(A_-)$ is recursively approximable. ■

As per the argument in Section 4.5, all of these results generalize to cases where α is not required to be constant outside $[-y^*, y^*]$. They hold for any computable α, provided $\alpha(1) < ½ < \alpha(-1)$, and for all $y \in [-1, 1]$, $0 < \alpha(y) < 1$.

Note that our function φ is non-computable because it is discontinuous. It might not be too difficult to replace φ with a continuous, computable function and still get the same results. Recall that the φ splits any horizontal into two segments, shifting one upwards and the other downwards. We might redefine φ so that in any horizontal spanning X, a small segment near $x = \alpha(y)$ is stretched out in some computable way by φ to connect the upward- and downward-shifted components. With suitable handling of the details, we could thus obtain non-d.m.z. long-term behavior by iteration of a computable map. (A modification of this kind was suggested by Wayne Myrvold.)

REFERENCES

Aberth, Oliver. 1980. *Computable Analysis.* New York: McGraw-Hill, Inc.

Alexander, J.C., James A. Yorke, Zhiping You, and Ittai Kan. 1992. Riddled basins. *International Journal of Bifurcation and Chaos* 2: 795-813.

Arnol'd, Vladimir I. 1963. Small denominators and problems of stability of motion in classical and celestial mechanics. *Russian Mathematical Surveys* 18: 85-191.

——. 1966. Instability of dynamical systems with several degrees of freedom.

Arnol'd, Vladimir I. and A. Avez. 1968. *Ergodic Problems of Classical Mechanics.* New York: W.A. Benjamin, Inc.

Barrow-Green, June. 1997. *Poincaré and the three-body problem.* American Mathematical Society/London Mathematical society.

Batterman, Robert W. 1998. Why equilibrium statistical mechanics works: Universality and the renormalization group. *Philosophy of Science* 65: 183-208.

Benacerraf, Paul, and Hilary Putnam, eds. 1983. *Philosophy of Mathematics: Selected Readings.* Cambridge: Cambridge University Press.

Bennett, Charles H. 1973. Logical reversibility of computation. *IBM Journal of Research and Development* 17: 525-532. Figure 2, p. 910, Copyright 1982 by Kluwer Academic Publishers, with kind permission of Springer Science and Media.

——. 1982. The thermodynamics of computation—a review. *International Journal of Theoretical Physics* 21: 905-940.

Berry, Arthur. [1898] 1961. *A Short History of Astronomy.* New York: Dover Publications, Inc.

Bishop, Errett. 1967. *Foundations of Constructive Analysis.* New York: McGraw-Hill.

Blum, Lenore, Felipe Cucker, Mike Shub, and Steve Smale. 1998. *Complexity and Real Computation.* New York: Springer-Verlag New York, Inc. Figure B, p. 40, Chapter 2, Copyright 1998 Springer-Verlag New York, Inc., with kind permission of Springer Science and Business Media.

Blum, Lenore, Mike Shub, and Steve Smale. 1989. On a theory of computation and complexity over the real numbers: NP-completeness, recursive functions, and universal machines. *Bulletin of the American Mathematical Society* 21: 1-46.

Blum, Lenore and Steve Smale. 1993. The Gödel incompleteness theorem and decidability over a ring. In M. W. Hirsch et al., eds., *From Topology to Computation: Proceedings of the Smalefest.* New York: Springer, 321-339.

Boldi, Paolo and Sebastiano Vigna. 1999. Equality is a jump. *Theoretical Computer Science* 219: 49-64.

Borel, Émile. 1912. Le calcul des intégrales definies. *Journal de Mathématiques Pures et Appliquées* 8: 159-210.

Boyd, Richard. 1972. Determinism, laws, and predictability in pinciple. *Philosophy of Science* 39: 431-450.

Brattka, Vasco. 2003a. Plottable real number functions. In Marc Daumas et al., eds., *RNC'5: Real Numbers and Computers.* Lyon: Institut National de Recherche en Informatique et en Automatique, 13-30.

——. 2003b. The emperor's new recursiveness: The epigraph of the exponential function in two models of computability. In Masami Ito and Tesuo Omaoka, eds., *Words, Languages, and Combinatorics* III. Rivers Edge: World Scientific Publishing.

Brattka, Vasco and Peter Hertling. 1996. Feasible real random access machines. In Keith G. Jeffrey, Jaroslav Král, and Miroslav Bartosek, eds., *SOFSEM'96: Theory and Practice of Informatics.* Vol. 1175 of *Lecture Notes in Computer Science.* Berlin: Springer, 335-342.

Brattka, Vasco and Klaus Weihrauch. 1999. Computability on subsets of Euclidean space I: closed and compact subsets. *Theoretical Computer Science* 219: 65-93.

Bridges, Douglas S. 1999. Constructive mathematics: A foundation for computable analysis. *Theoretical Computer Science* 219: 95-109.

Brouwer, L. E. J. [1912] 1983. Intuitionism and formalism. *Bulletin of the American Mathematical Society* 20, 1913. Translated by Arnold Dresden. Reprinted in Benacerraf and Putnam 1983, 77-89.

Browder, Felix E., ed. 1976. *Mathematical Developments Arising from Hilbert Problems*. Proceedings of Symposia in Pure Mathematics, vol. 28. American Mathematical Society.

Campagnolo, Manuel L., Cristopher Moore, and José F. Costa. 1998. Analog computers and the iteration functional. Technical report DI-FCUL TR-98-10. Departamento de Informática, Faculdade de Ciências da Universidade de Lisboa. http://www.di.fc.ul.pt/biblioteca/tech-reports.

Celletti, Alessandra, and Luigi Chierchia. 1997. On the stability of realistic three-body problems. *Communications in Mathematical Physics* 186, 413-449.

Chierchia, L., and G. Gallavotti. 1982. Smooth prime integrals for quasi-integrable Hamiltonians. *Il Nuovo Cimento* 67 B, 277-295.

Chirikov, Boris V. 1979. A universal instability of many-dimensional oscillator systems. *Physics Reports* 52, 263-379.

Church, Alonzo. [1936] 1965. An unsolvable problem of elementary number theory. *American Journal of Mathematics* 58: 346-363. Reprinted in Davis 1965, 89-107.

da Costa, N.C.A., and F.A. Doria. 1993. On Arnol'd's Hilbert Symposium problems. In Georg Gottlob et al., eds., *Computational Logic and Proof Theory: Third Kurt Gödel Colloquium, KGC '93*. Berlin: Springer-Verlag.

Davis, Martin, ed. 1965. *The Undecidable.* Hewlett, New York: Raven Press.

Devaney, Robert L. 1989. *An Introduction to Chaotic Dynamical Systems*. Redwood City, California: Addison-Wesley Publishing Company, Inc.

Diacu, Florin. 1996. The solution of the *n*-body probem. *Mathematicla Intelligencer* 18: 66-70.

Diacu, Florin, and Philip Holmes. 1996. *Celestial Encounters: The Origins of Chaos and Stability.* Princeton: Princeton University Press.

Dugundji, James. 1966. *Topology*. Boston: Allyn and Bacon, Inc.

Earman, John. 1986. *A Primer on Determinism.* Dordrecht: D. Reidel Publishing Company.

Edmonds, David, and John Eidinow. 2001. *Wittgenstein's Poker: The Story of a Ten-Minute Argument Between Two Great Philosophers.* New York: HarperCollins Publishers, Inc.

Feynman, Richard P. 1986. Quantum mechanical computers. *Foundations of Physics* 16: 507-531.

Fine, Arthur. [1986] 1996. *The Shaky Game.* Chicago: University of Chicago Press.

Feudel, Ulrike, and Celso Grebogi. 1997. Multistability and the control of complexity. *Chaos* 7: 597-604.

Folland, Gerald B. 1984. *Real Analysis.* New York: John Wiley & Sons.

Fredkin, Edward and Tommaso Toffoli. 1982. Conservative logic. *International Journal of Theoretical Physics* 21: 219-253.

Gödel, Kurt. [1931] 1986. Über formal unenterscheidbare städer *Principia Mathematica* und verwandter systeme I. *Monatshefte für Mathematik und Physik* 38: 173-198. Translated in Gödel 1986, 145-195.

—. [1931a] 1986. Discussion on the Foundations of Mathematics. Excerpt from Hans Hahn et al., Diskussion zur Grundlegung der Mathematik, *Erkenntnis* 2: 135-51. Translated by John W. Dawson in Gödel 1986, 201-205.

—. [1934] 1965. On undecidable propositions of formal mathematical systems. Notes from lectures at the Institute for Advanced Study, Feb. to May 1934, taken by S. C. Kleene and B. Rosser. In Davis 1965, 41-74. Reprinted in Gödel 1986, 346-371.

—. 1986. *Collected Works* I. Oxford: Oxford University Press.

Goroff, Daniel L. 1993. Introduction to Poincaré [1892-1899] 1993.

Grzegorczyk, Andzrej. 1955. Computable functionals. *Fundamenta Mathematicae* 42: 169-202.

—. 1957. On the definitions of computable real continuous functions. *Fundamenta Mathematicae* 44: 61-67.

Heagy, James F., Thomas L. Carroll, and Louis M. Pecora. 1994. Experimental and numerical evidence for riddled basins in coupled chaotic systems. *Physical Review Letters* 73: 3528-3531.

Hemmerling, Armin. 2003. Approximate decidability in Euclidean spaces. *Mathematical Logic Quarterly* 49: 34-56.

Herbrand, Jacques. 1931] 1967. On the consistency of arithmetic. In van Heijenoort 1967, 618-628.

Herman, Michael. 1998. Some open problems in dynamical systems. *Documenta Mathematica*, Extra Volume ICM 1998 II: 797-808. (http://www.emis.de/journals/DMJDMV/xvol-icm/09/09.html)

Hertling, Peter. 1999. An effective Riemann mapping theorem. *Theoretical Computer Science* 219: 225-265.

Heyting, Arend. [1956] 1983. Disputation. From A. Heyting, *Intuitionsm, an Introduction,* North-Holland, 1956. Reprinted in Benacerraf and Putnam 1983, 66-76.

Hilbert, David. [1926] 1983. On the infinite. From *Mathematische Annalen* 95: 161-190. Translated by Erna Putnam and Gerald J. Massey, in Benacerraf and Putnam 1983, 183-201.

Hoel, Paul G., Sidney C. Port, and Charles J. Stone. 1971. *Introduction to Probability Theory.* Boston: Houghton Mifflin Company.

Kan, Ittai. 1994. Open sets of diffeomorphisms having two attractors, each with an everywhere dense basin. *Bulletin (New Series) of the American Mathematical Society* 31: 68-74.

Kleene, Stephen C. 1986. Introductory note to [Gödel] *1930b, 1931* and *1932b.* In Gödel 1986, 126-140.

——. 1986a. Introductory note to [Gödel] *1934.* In Gödel 1986, 338-345.

Ko, Ker-I. 1991. *Complexity Theory of Real Functions.* Boston: Birkhauser.

Ko, Ker-I, and H. Friedman. 1982. Computational complexity of real functions. *Theoretical Computer Science* 20: 323-352.

Kreitz, C., and Klaus Weihrauch. [1984] 1987. Compactness in Constructive Analysis Revisited. Informatik-Berichte, Fernuniversität Hagen, 1984. Also in *Annals of Pure and Applied Logic* 36 (1987): 29-38.

Kriesel, Georg. [1958] 1983. Hilbert's programme. *Dialectica* 12: 346-372. Reprinted in Benacerraf and Putnam 1983, 207-240.

Kummer, Martin, and Marcus Schäefer. 1995. Computability of convex sets. In *Proceedings of STACS 1995.* Vol. 900 of *Lecture notes in Computer Science.* Berlin: Springer: 550-561.

Lacombe, Daniel. 1957. Les ensembles récursivement ouverts ou fermés, et leurs applications à l'analyse récursive. *Comptes Rendus de l'Académie des Sciences Paris* 245: 1040-1043.

——. 1958. Les ensembles récursivement ouverts ou fermés, et leurs applications à l'analyse récursive. *Comptes Rendus de l'Académie des Sciences Paris* 246: 28-31.

Lakatos, Imre. [1963] 1976. *Proofs and Refutations.* Cambridge: Cambridge University Press.

——. 1970. Falsification and the methodology of scientific research programmes. In Lakatos and Alan Musgrave, eds., *Criticism and the Growth of Knowledge,* Cambridge: Cambridge University Press, 91-196.

Laplace, Pierre Simon. [1814] 1951. *A Philosophical Essay on Probabilities.* New York: Dover Publications, Inc.

——. 1878-1912. *Oeuvres Complètes.* Paris: Gauthier-Villars.

Leibniz, Gottfried Wilhelm. [1715] 1989. Leibniz's first paper, being an extract of a letter. In *Philosophical Essays.* Indianapolis: Hackett Publishing Co., 320-321.

Malament, David, and Sandy Zabell. 1980. Why Gibbs phase averages work—The role of ergodic theory. *Philosophy of Science* 47: 339-349.

Mazur, S. 1963. Computable Analysis. *Razprawy Matematyczne* 33: 1-111.

Milnor, John. 1985. On the concept of attractor. *Communications in Mathematical Physics* 99: 177-195.

Moore, Cristopher. 1990. Unpredictability and undecidability in dynamical systems. *Physical Review Letters* 64: 2354-2357.

——. 1991. Generalized shifts: Unpredictability and undecidability in dynamical systems. *Nonlinearity* 4: 199-230.

Moser, Jürgen. 1973. *Stable and Random Motions in Dynamical Systems.* Princeton, New Jersey: Princeton University Press and University of Tokyo Press.

——. 1978. Is the solar system stable? *Mathematical Intelligencer* 1: 65-71.

——. 1998. Dynamical systems—past and present. *Documenta Mathematica*, Extra Volume ICM 1998 I: 381-402. http://www.emis.de/journals/DMJDMV/xvol-icm/00/00.html

Moschovakis, Yiannis. 1980. *Descriptive Set Theory.* Amsterdam: North-Holland Publishing Company.

Myrvold, Wayne C. 1994. Constructivism, computability, and physical theories. Ph.D. dissertation, Boston University.

—. 1995. Computability in quantum mechanics. In Werner Depauli-Schimanovich, Eckhart Köhler, and Friedrich Stadler, eds., *The Foundational Debate: Complexity and Constructivity in Mathematics and Physics.* Dordrecht: Kluwer Academic Publishers, 33-64.

—. 1997. The decision problem for entanglement. In R. S. Cohen et al., eds., *Potentiality, Entanglement, and Passion-at-a-Distance.* Great Britain: Kluwer Academic Publishers, 177-190.

Newton, Isaac. [1687] 1934. *Principia*, Volume I. Berkeley, California: University of California Press.

—. [1715] 1979. *Opticks.* New York: Dover Publications.

Ott, Edward, J. C. Alexander, Ittai Kan, John C. Sommerer, and James A. Yorke. 1994. The transition to chaotic attractors with riddled basins. *Physica* D 76: 384.

Ott, Edward, and John C. Sommerer. 1994. Blowout bifurcations: the occurrence of riddled basins and on-off intermittency. *Physics Letters* A 188: 39-47.

Parker, Matthew W. 2003. Undecidability in $\mathbf{R}^n$: Riddled basins, the KAM tori, and the stability of the solar system. *Philosophy of Science* 70: 359-382.

Peirce, Charles Sanders. 1878. How to make our ideas clear. *Popular Science Monthly* 12: 286-302.

Penrose, Roger. 1989. *The Emperor's New Mind: Concerning Computers, Minds, and the Laws of Physics.* New York: Oxford University Press.

Pitowsky, Itamar. 1996. Laplace's demon consults an oracle: The computational complexity of prediction. *Studies in the History and Philosophy of Modern Physics* 27: 161-180. Figure reprinted from same, p. 166, Copyright 1996 Elsevier Science Ltd., with permission from Elsevier.

Plato, Jan von. 1983. The method of arbitrary functions. *British Journal for the Philosophy of Science* 34: 37-47.

Poincaré, Henri. 1890. Sur le problème des trois corps et les équations de la dynamique. *Acta Mathematica* 13: 1-270.

——. 1891. Le problème des trois corps. *Revue Générale des Sciences Pures et Appliquées* 2 (1): 1-5.

——. [1892-1899] 1957. *Les Méthodes Nouvelles de la Méchanique Céleste.* Reprint. New York: Dover Publications, Inc.

——. [1892-1899] 1993. *New Methods of Celestial Mechanics.* Translation of Poincaré [1892-1899] 1957. Woodbury, New York: American Institute of Physics.

——. 1898. On the stability of the solar system. *Nature* 58: 183-185.

——. [1907] 1952. Le hasard. *La Revue du Mois* 3 (1907): 257-276. Translated in Henri Poincaré 1952, *Science and Method,* New York: Dover Publications, Inc.

——. [1908] 1952. *Science and Method.* New York: Dover Publications, Inc.

Popper, Karl. [1934] 1968. *Logik der Forschung.* Translated as *The Logic of Scientific Discovery.* Harper and Row.

Pöschel, Jürgen. 2001. A lecture on the classical KAM theorem. *Proceedings of Symposia in Pure Mathematics* 69: 707-732.

Pour-El, Marian Boykan, and Ian Richards. 1981. The wave equation with computable initial data such that its unique solution is not computable. *Advances in Mathematics* 39: 215-239.

——. 1983. Computability and noncomputability in classical analysis. *Trans. Am. Math. Soc.* 275: 539-545.

Preparata, F. P., and M. I. Shamos. 1985. *Computational Geometry.* Berlin: Springer-Verlag.

Russell, Bertrand. [1912-1913] 1963. On the notion of cause. *Proceedings of the Aristotelian Society* 8. Reprinted in *Mysticism and Logic and Other Essays.* New York: Barnes and Noble, 132-151.

Saari, Donald G. 1971-1973. Improbability of collisions in Newtonian gravitational systems. Part 1, Erratum, and Part 2. *Transactions of the American Mathematical Society* 162: 267-271; 168: 571; 181: 351-368.

Šanin, N. 1968. *Constructive Real Numbers and Function Spaces.* Translations of Mathematical Monographs 21, transl. by E. Mendelson. Providence, Rhode Island: American Mathematical Society.

Sartre, Jean-Paul. 1948. *Existentialism and Humanism.* London: Methuen & Co., Ltd.

Shepherdson, J. C., and H. E. Sturgis. 1963. Computability of recursive functions. *Journal of the Association for Computing Machines* 10: 217-255.

Siegelmann, Hava T. 1996. The simple dynamics of super Turing theories. *Theoretical Computer Science* 168: 461-472.

Sklar, Lawrence. 1973. Statistical explanation and ergodic theory. *Philosophy of Science* 40: 194-212.

—. 1993. *Physics and Chance: Philosophical Issues in the Foundations of Statistical Mechanics*. Cambridge, England: Cambridge University Press.

Skolem, Thoralf. [1923] 1967. The foundations of elementary arithmetic established by means of the recursive mode of thought, without the use of apparent variables ranging over infinite domains. In van Heijenoort 1967, 302-333.

Smith, Peter. 1998. *Explaining Chaos*. Cambridge, England: Cambridge University Press.

Soare, Robert I. 1987. *Recursively Enumerable Sets and Degrees*. New York: Springer-Verlag.

Sommerer, John C., and Edward Ott. 1993. A physical system with qualitatively uncertain dynamics. *Nature* 365: 138-140. Figure reprinted from same, Copyright 1996 Elsevier Science Ltd., with permission from Elsevier.

—. 1996. Intermingled basins of attraction: Uncomputability in a simple physical system. *Physics Letters* A 214: 243-251.

Turing, Alan. [1936-1937] 1965. On computable numbers, with an application to the entscheidungsproblem. *Proceedings of the London Mathematical Society*. Series 2. 42: 230-265. Errata in v. 43: 544-546. Reprinted in Davis 1965, 116-154.

van Heijenoort, Jean, ed. 1967. *From Frege to Gödel: A Sourcebook in Mathematical Logic*. Cambridge, Massachusetts: Harvard University Press.

Vranas, Peter B. M. 1998. Epsilon-ergodicity and the success of equilibrium statistical mechanics. *Philosophy of Science* 65: 688-708.

Wang Qiu-dong. 1991. The global solution of the n-body problem. *Celestial Mechanics and Dynamical Astronomy* 50: 73-88.

Weihrauch, Klaus. 1985. Type 2 recursion theory. *Theoretical Computer Science* 38: 17-33.

——. 1993. Computability on computable metric spaces. *Theoretical Computer Science* 113: 191-210.

——. 2000. *Computable Analysis: An Introduction.* Berlin: Springer. Figure 5.2, p. 127, Chapter 5, Copyright 2000 Springer-Verlag Berlin Heidelberg, with kind permission of Springer Science and Business Media.

Wittgenstein, Ludwig. [1933-1934] 1964. *The Blue and Brown Books.* New York: Harper and Row.

——. 1953. *Philosophical Investigations.* New York: The MacMillan Company.

Wolfram, Stephen. 1985. Undecidability and intractability in theoretical physics. *Physical Review Letters* 54: 735-738.

——. 2002. *A New Kind of Science.* Champaign, IL: Wolfram Media, Inc.

Wylie, Ray C., and Louis C. Barrett. 1995. *Advanced Engineering Mathematics.* New York: McGraw-Hill, Inc.

Ziegler, Martin. 2002. Computability on regular subsets of Euclidean space. *Mathematical Logic Quarterly* 48, Supplement 1: 157-181.

Zhou, Qing. 1996. Computable real-valued functions on recursive open and closed subsets of Euclidean space. *Mathematical Logic Quarterly* 42: 379-409.