BROWNIAN MOTION ON NESTED FRACTALS

by

Tom Lindstrøm

Department of Mathematics

University of Oslo, Norway

ABSTRACT

The purpose of this paper is to construct Brownian motion on a reasonably general class of self-similar fractals. To this end, I introduce an axiomatically defined class of "nested fractals", which satisfy certain symmetry and connectivity conditions, and which also are (in the physicists' terminology) finitely ramified. On each one of these nested fractals, a Brownian motion is constructed and shown to be a strong Markov process with continuous paths. If the Laplacian Δ on the fractal is defined as the infinitesimal generator of the Brownian motion, and $n(\alpha)$ denotes the number of eigenvalues of $-\Delta$ less than α, I prove that

$$n(\alpha) \sim \alpha^{d \cdot \log\nu/\log\lambda} \quad \text{as} \quad \alpha \to \infty,$$

where d is the Hausdorff dimension of the fractal, and ν and λ are two parameters describing its self-similarity structure. In general, $\log\nu/\log\lambda \neq 1/2$ and hence Weyl's conjecture can only hold for fractals in a modified form.

AMS 1980 subject classifications: Primary 60J60, 60J65,
secondary 03H05, 35P20

Key words and phrases: Brownian motion on fractals, diffusions, self-similar fractals, fractal Laplacians, asymptotic distribution of eigenvalues, nonstandard analysis.

TABLE OF CONTENTS

I. Introduction

In the last few years there has been an increased interest among mathematicians for fractals and fractal media. Much of the motivation comes from physics; as physicists unearth ever more convincing examples of fractal structure in nature, more and more mathematicians are persuaded that the theory of fractals is not a collection of exotic and useless counterexamples, but a promising and exciting tool for describing and understanding important physical phenomena. Nobody could have argued this case more vividly and flamboyantly that Benoit Mandelbrot in the various editions of his books [25], [26], [27].

One of many ways in which physicists run into fractals is by observing substances with a noninteger dimension (in a sense I shall describe in the next chapter). Natural mathematical models for such substances are the sets of noninteger Hausdorff dimension that have been known to mathematicians for more than half a century. But these sets in themselves do not give us much information about the important physical properties of fractal media such as charge distribution, heat transfer, wave propagation, and quantum effects. In fact, since the Laplace operator is the crucial ingredient in the equations governing all these phenomena (the potential equation, the heat equation, the wave equation, and the Schrödinger equation), one might argue that a reasonable model for fractal media necessarily includes a description of the associated Laplacians (assuming, of course, that a fractal has a "natural" Laplace operator more or less the same way a Riemannian manifold has a "natural" Laplace-Beltrami operator).

Now there is an obvious obstacle to constructing Laplacians on fractals: Laplacians are differential operators and fractals do

not have a differentiable structure - they are the classical examples of nondifferentiable objects. For the probabilist there is an easy way to circumvent this problem; to him or her the Laplacian is just the infinitesimal generator of Brownian motion, and to describe Brownian motion one does not need a differentiable structure.

This observation will be my point of departure in the present paper. Starting with an axiomatically defined subclass of Hutchinson's [18] self-similar fractals, I'll show that each one of these "nested fractals" (as I call them) supports a Brownian motion in a natural way (- for all I know, it may actually support more than one Brownian motion; this is one of the big open questions in the theory). The infinitesimal generator of such a process will be a self-adjoint operator on $L^2(E,m)$ (where E is the fractal and m is the invariant measure on E), and I'll refer to these operators as Laplacians on fractals. They turn out to have some rather unexpected properties; e.g., if

$$n(\alpha) = \text{the number of eigenvalues of } -\Delta \text{ less than } \alpha,$$

it's natural to conjecture (see e.g. [6], [7]) that Weyl's classical law for the asymptotic distribution of eigenvalues will generalize to

$$n(\alpha) \sim \alpha^{d/2} \qquad \text{as } \alpha \to \infty$$

where d is the Hausdorff dimension of the fractal, but this turns out to be false. Instead, there is an asymptotic expression

$$n(\alpha) \sim \alpha^{d \cdot \log \nu / \log \lambda} \qquad \text{as } \alpha \to \infty, \tag{1.1}$$

where ν and λ are two parameters describing the fractal. This result is not quite as surprising as it may seem at first glance; it's a consequence of the simple observation that in the fractal case the relationship between the time increment Δt and the space

increment Δx of Brownian motion is not the customary $\Delta x \sim \Delta t^{1/2}$ but $\Delta x \sim \Delta t^{\log\nu/\log\lambda}$. For the Sierpinski gasket (see Example II.1), formula (1.1) has already been proved by Barlow and Perkins [5], and without formal proof (but with various heuristic arguments) it has been known in the physics literature for some time (see [30]).

The paper is organized as follows: In the next chapter, I discuss informally the relationship between the physical and the mathematical approach to fractals. You may find this an unduly long piece of expository writing for what is basically a research paper, but to my mind it serves two important purposes; it explains the physical relevance of fractals and Hausdorff dimension (which is not always easy to understand from the literature), and it provides the reader with the necessary examples to better appreciate the axiomatic theory I develop later in the paper. In the same expository style, I use Chapter III to describe the fundamental ideas of the paper on a simple but typical example; again I hope that this informal introduction will help the reader to see through the abstract machinery needed for the general case. The systematic development of the theory begins in Chapter IV where I define axiomatically the class of "nested fractals" that I'm going to work with. The most important and also the most restrictive of my postulates is the Nesting Axiom which - in the physicists' terminology - says that a nested fractal is "finitely ramified". After the axioms have been introduced, the rest of Chapter IV is devoted to a systematic account of their basic consequences.

Brownian motion on a nested fractal will be constructed as the limit of a sequence of random walks, and to avoid difficult combinatorial problems, it's important to choose the transition probabilities of these random walks as cleverly as possible. In Chapter V, I explain how this can be done with an appeal to Lefschetz's fixed point theorem. If Brownian motion is to be obtained

as a limit of random walks B_n, I need to know how to rescale time as n goes to infinity, and using the Perron-Frobenius theory of positive matrices, this question is dealt with in Chapter VI. In Chapter VII, I can finally construct Brownian motions on nested fractals and show that they are Feller-Dynkin processes with continuous paths. As already said, I originally choose the transition probabilities of my random walks with extreme care, but in Chapter VIII I prove an invariance principle which shows that all this care is not really necessary; under certain conditions we always get the same limit process no matter which transition probabilities we start with. In Chapter IX, I define Laplacians on fractals as infinitesimal generators of Brownian motions; show that they are self-adjoint operators; and derive formula (1.1) for the asymptotic distribution of their eigenvalues. The tenth and last chapter contains a list of open problems and a few remarks.

I should perhaps say a few words about one rather untraditional aspect of my presentation - the use of nonstandard analysis. Before I scare anybody away, let me say that it doesn't play a very central part in this paper; with the exception of a guest appearance in the proof of Lemma VI.4, it doesn't enter the scene until the beginning of Chapter 7, and by then all the main ideas have been presented, and what remains is to run the technical machinery. Since the technical machinery of nonstandard analysis is ideally suited for the problems at hand, I have decided to use it rather than, say, the theory of weak convergence. The readers who want to come along can find introductions to nonstandard probability theory in [1] and [33], and the readers who instead want to rephrase the results in terms of their own favorite technique, will, I'm convinced, not find this very difficult (- the crash course in nonstandard analysis provided by the ten first pages of [23] may be helpful).

Let me finally say a few words about the relationship between the present paper and other contributions to the literature, both the physical and the mathematical. In both fields the attention concentrates almost exclusively on Brownian motion on the Sierpinski gasket. A good introduction to the physical theory is Rammal and Toulouse [30]; other papers of interest are [3], [17], [29]. On the mathematical side, the papers by Kusuoka [21], Goldstein [16], and Barlow and Perkins [5] all contain rigorous constructions of Brownian motion on the Sierpinski gasket. The first two of these papers were known to me when I started to write the present paper, the last one only appeared when I was almost finished (but I did know about the main results through private communication). In a certain sense, the present paper is complementary to the previous ones as it concentrates mainly on problems that do not exist for the Sierpinski gasket, but there are overlaps, of course, especially with Barlow's and Perkins' work; both they and I estimate the modulus of continuity for the Brownian sample path and find the asymptotic distribution of eigenvalues for the Laplacian (in either case they do it better than I do, but only for the Sierpinski gasket). It's interesting to note that when we attack the same problem, we usually do it by completely different methods, and thus the papers are complementary also in this respect. The only place where I have actually used ideas from Barlow's and Perkins' paper is in estimating the Green functions in Chapter IX, but I acknowledge their priority to any result the two papers have in common. All the contributions mentioned so far are basically of a probabilistic nature, but a more potential theoretic approach to the same bundle of problems is presented in the recent paper by Metz [28].

An exception to the rule that everybody works with the Sierpinski gasket, is the recent paper by Barlow and Bass [4] on Brownian motion on the Sierpinski carpet (see Example IV.6). As far as I know, this is the only mathematical construction of a Brownian motion on a (nontrivial) fractal which is not covered by the theory in the present paper. The Sierpinski carpet does not satisfy the most restrictive of my axioms (the Nesting Axiom), and requires a totally different approach than the one used in this paper. A good understanding of Brownian motion on the Sierpinski carpet would mean a great step forward toward a general theory for diffusions on fractals, but Barlow's and Bass' paper indicates that we are here faced by very deep and difficult problems. At this point, let me just mention my own earlier and very sketchy contribution [24], which isn't worth much as it stands today, but which contains ideas I hope some day will become useful.

I would also like to mention three interesting papers on different, but related subjects; one is Kesten's paper [20] on random walks on percolation clusters (see also [30] for the physical background), the other two are the papers by Brossard and Carmona [8] and Lapidus and Fleckinger-Pellé [22] on the asymptotic distribution of eigenvalues for the Laplacian on domains with fractal boundary.

Acknowledgements: I'm grateful to Martin Barlow, Ed Perkins and Volker Metz for their generosity in sharing their results with me before they were published. I would also like to thank Sergio Albeverio, Ward Henson and Nils Øvrelid for helpful conversations and useful references to the literature. Finally, many thanks to Tove Lieberg who typed the paper with skill and efficiency.

II. Fractals in physics and mathematics.

As I mentioned in the Introduction, one way physicists run into fractals is by observing objects with a noninteger dimension. To understand what this means, let us assume that we are given a three-dimensional, homogeneous solid, and that we measure the mass of that part of the solid that is inside a cubic box of side ℓ for different values of ℓ. What we find is, of course, that the mass $M(\ell)$ is proportional to ℓ^3. If we do the same experiment with a two-dimensional or a one-dimensional object (still using three-dimensional boxes), we find that $M(\ell)$ is roughly proportional to ℓ^2 in the first case and to ℓ in the second case. In all three cases we recover the dimension d as the exponent of ℓ; we have

$$M(\ell) \approx C\ell^d$$

or, taking logarithms,

$$\log M(\ell) \approx d \log \ell + C'$$

(I'm using $\approx$ to mean "approximately equal" in an informal sense). The last formula suggests how an intelligent creature with no intuitive sense of dimension could try to measure the dimension of various objects by carrying out the experiment above and plotting the results in a double-logarithmic coordinate system as shown in Figure 1.

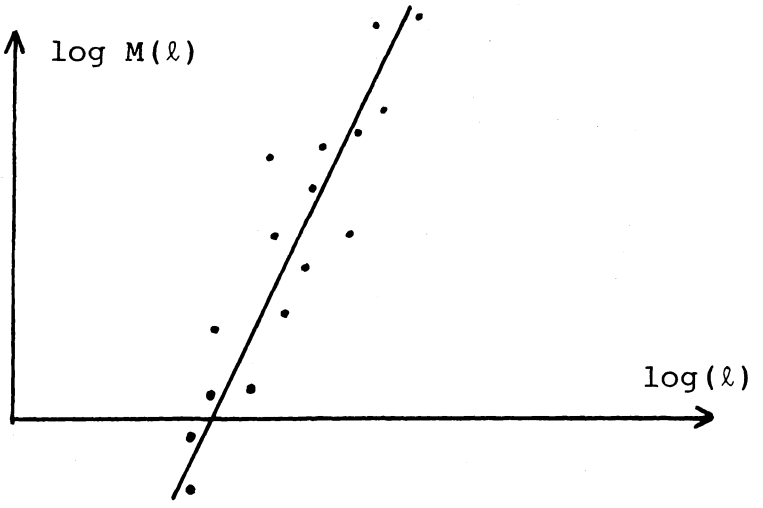


Figure 1

The slope of the best fitting, straight line through the data is then the experimental value for the dimension.

Experiments of this kind are much more than mere thought experiments; physicists have actually carried them out for a wide variety of different objects. If the object has an irregular boundary and contains holes of all sizes, they often find a sharply defined experimental value d which is not an integer, and hence they conclude that the object is fractal in the sense that it has a noninteger dimension. Examples are gases or liquids injected in porous media (e.g. oil or water in rock), aggregations of particles and colloids, and surfaces of turbulent clouds, just to mention a few. The recent book by Feder [13] contains a detailed description of these phenomena and many more, and is warmly recommended to anybody who wants to understand the rôle of fractal media in solid state physics. As will be seen from Feder's book, the experiment outlined above is only one of many actually used by physicists to measure dimension; the other ones are based on similar, intuitive scaling properties of sets of different dimensionality.

One must, of course, be a little careful in interpreting the results of such experiments; it's obvious, for instance, that the linear relationship between $\log M(\ell)$ and $\log \ell$ in the experiment above will break down when ℓ becomes too small (the size of an atom, if not before). But there is nothing new in this; if we carry out the experiment for a real, physical surface, the relationship $M(\ell) \sim \ell^2$ will only hold as long as ℓ is much larger than the thickness of the surface: we can only treat a surface as a two-dimensional object as long as we stick to the proper physical scale, and the same principle applies to fractal media.

Let us now compare these physical examples of fractal behavior to the fractals created by mathematicians. The first example is the wellknown Sierpinski gasket.

II.1 Example. We begin with the isosceles triangle T_0 in Figure 2a).

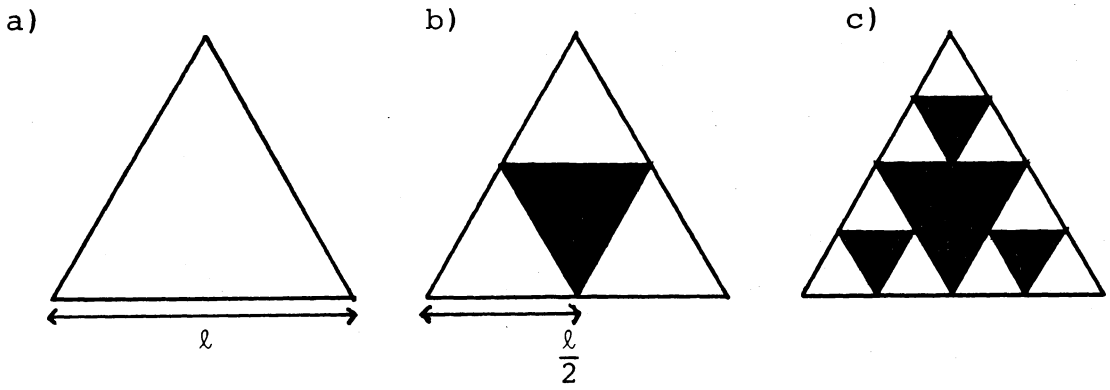


Figure 2

Removing the black triangle in the middle of Figure 2b), we are left with three triangles similar to but smaller than the original one. Repeating the procedure with each one of these triangles, we get the nine triangles in Figure 2c). Again we repeat the procedure with each one of these triangles, and so on to infinity. The limit set T is the Sierpinski gasket.

Let us try to estimate the dimension of T by means of the physicists' method described above. We return for a moment to Figure 2b), but imagine that the construction has been carried through to the end so that each of the three white regions is full of triangular holes of all sizes. In an intuitive sense each one of these white triangles has size one third of the whole set, while the length of its side is one half of the original side.

Hence

$$M(\frac{\ell}{2}) = \frac{1}{3} M(\ell)$$

(where $M(\alpha)$ - in our obvious, intuitive sense - measures the "size" or "mass" of a triangular region with side α). Repeating the argument, we see that

$$M(\frac{\ell}{2^n}) = \frac{1}{3^n} M(\ell)$$

for all positive integers n. This shows that

$$M(\lambda) = C\lambda^{\log 3/\log 2}$$

at least when λ is of the form $\ell/2^n$, indicating strongly that T has dimension $\log 3/\log 2$. And indeed, it's not hard to prove by strict mathematical means (using, e.g., a theorem by Hutchinson [18] which I will state as Theorem IV.4) that the Hausdorff dimension of T is exactly $\log 3/\log 2$.

What this example shows is that at least some of the fractals constructed by mathematicians have scaling properties similar to those physicists find in nature, and that their physical "scaling dimension" often agrees with the mathematical notion of Hausdorff dimension. To give a little more substance to this claim, I'll sketch quickly two more examples of the same kind. The first is the Koch-curve which in many respects is simpler than the Sierpinski gasket, although it's construction is slightly more complicated in the sense that it is not only a question of cutting away unwanted bits and pieces from the original figure, but of actually changing the contour of the domain.

II.2 Example. Beginning with the line segment of length ℓ in Figure 3a), I first replace it by the four line segments of length $\ell/3$ in Figure 3b), and then I replace each one of these by four

segments of length $\ell/9$ as in Figure 3c), etc. The limit set is the Koch-curve.

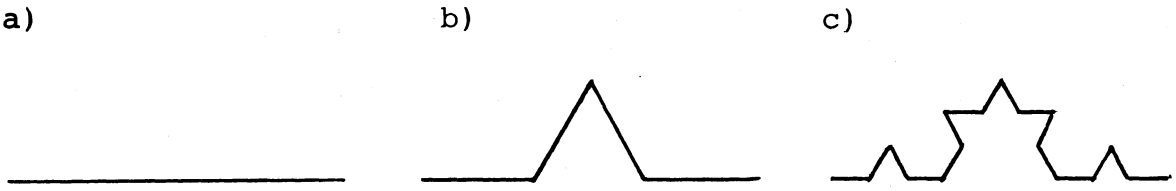


Figure 3

Estimating the dimension heuristically as in the previous example, I get

$$M(\frac{\ell}{3^n}) = \frac{1}{4^n} M(\ell),$$

indicating that $M(\ell) \sim \ell^{\log 4/\log 3}$ and $d=\log 4/\log 3$. Again it is easy to use Hutchinson's general result to check that the Hausdorff dimension of the Koch-curve is indeed $\log 4/\log 3$.

I claimed above that one of the reasons why physicists got interested in fractals was because they measure sets with noninteger dimensions. The next example shows that even sets with an integer dimension may have all the features we usually associate with fractals.

II.3 Example. (Nested cubes) We begin with the cube of side ℓ in Figure 4a), and replace it by the nine cubes of side $\ell/3$ in Figure 4b), etc. Estimating the dimension as before, we get

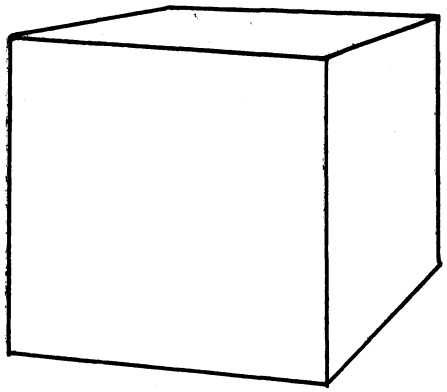
$$M(\frac{\ell}{3^n}) = \frac{1}{9^n} M(\ell),$$

which yields

$$M(\ell) \sim \ell^{\log 9/\log 3}$$

and $d = \log 9/ \log 3 = 2$!

a)



b)

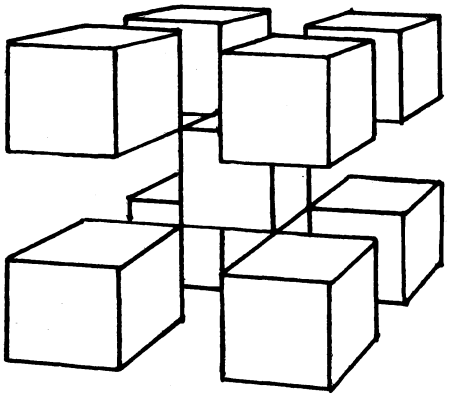


Figure 4

Also in this case the heuristic calculation is easily verified; the limit set of nested cubes is indeed two-dimensional. In addition to showing that fractal sets may very well have an integer dimension, this example also indicates that knowing the Hausdorff dimension doesn't tell us all that much about the set; the plane and the fractal in this example both have dimension 2, but they are obviously quite different in many other respects.

The limit sets in the three examples above all belong to the class of "nested fractals" which I will introduce in Chapter IV, and it may be worthwhile at this point to take a brief look at what they have in common. In each case an original geometric figure is replaced by a finite number of similar but smaller copies, and this procedure is then iterated an infinite number of times. To a large extent the construction can be described in terms of two parameters, the volume scaling factor μ and the linear scaling factor ν. The volume scaling factor μ is just the number of copies the original figure is replaced by at the first stage of the construction, while the linear scaling factor ν measures how many times larger the diameter of the original figure is than the diameter of each one of the copies. In the three examples above we have $\mu=3$, $\nu=2$ for the Sierpinski gasket; $\mu=4$, $\nu=3$ for the Koch-curve; and $\mu=9$, $\nu=3$ for the nested cubes. Note that in each case the Hausdorff dimension is

$$d = \frac{\log \mu}{\log \nu} ,$$

and, in fact, the general theorem by Hutchinson [18] which I have been referring to repeatedly, says that this will always be the case provided the copies (of all generations) are kept from overlapping too much (see Theorem IV.4).

III. Brownian motion on a snowflake.

In this chapter, I'll try to explain the basic ideas of the paper on an example which is sufficiently typical to exhibit all the conceptual difficulties of the general theory, but which at the same time is so simple geometrically that I can avoid most of the abstract machinery needed for the axiomatic approach. The example I have in mind is the snowflake fractal S obtained by the procedure illustrated in Figure 5.

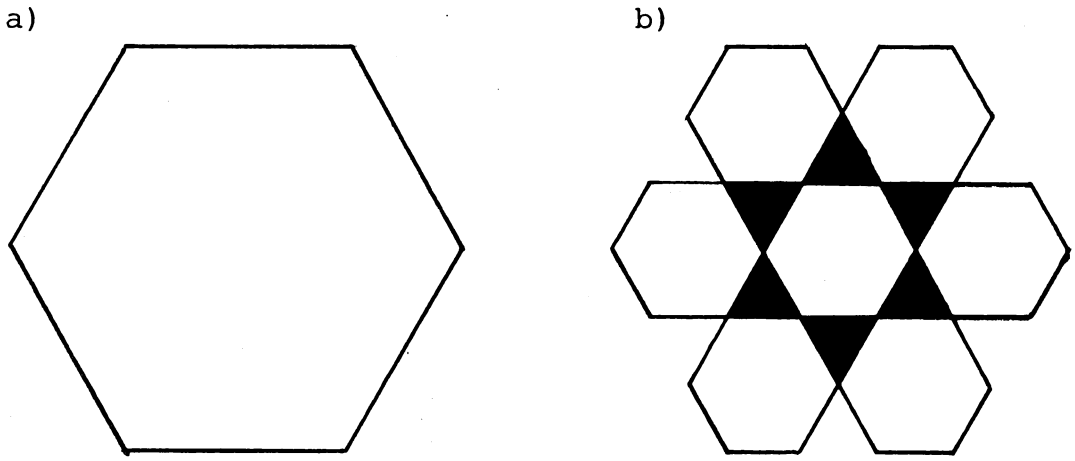


Figure 5

Starting with the regular hexagon S_0 in a), I replace it by the seven smaller hexagons in b), etc. The figure I have at stage n of the construction is called S_n, and S itself is just the limit of S_n as n goes to infinity. In the terminology of the last chapter, the volume scaling factor μ is 7, and the linear scaling factor ν is 3. As usual, one can show that the Hausdorff dimension of S is given by

$$d = \frac{\log \mu}{\log \nu} = \frac{\log 7}{\log 3}$$

I shall construct a Brownian motion B on S as the limit of random walks on S_n as n tends to infinity. Assume that Figure 6 shows a few of the 7^n hexagons of side 3^{-n} making up S_n.

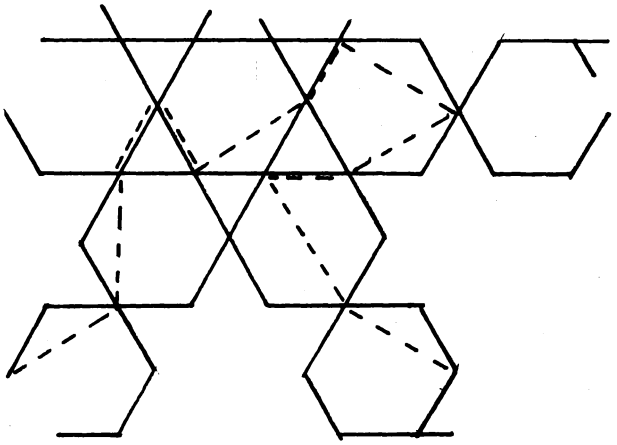


Figure 6

Two distinct vertices are called <u>neighbors</u> if they belong to the same one of these minimal hexagons. A <u>walk</u> on S_n is just a sequence $v_1, v_2, \ldots, v_p$ of vertices such that v_i and v_{i+1} are always (distinct) neighbors. The dotted path in Figure 6 connects consecutive elements in a walk on S_n.

To obtain a probability measure on the set of all possible walks, it's natural to assign a transition probability to each individual transition. I shall do this in the following way. Assume that we are in state 0 in Figure 7.

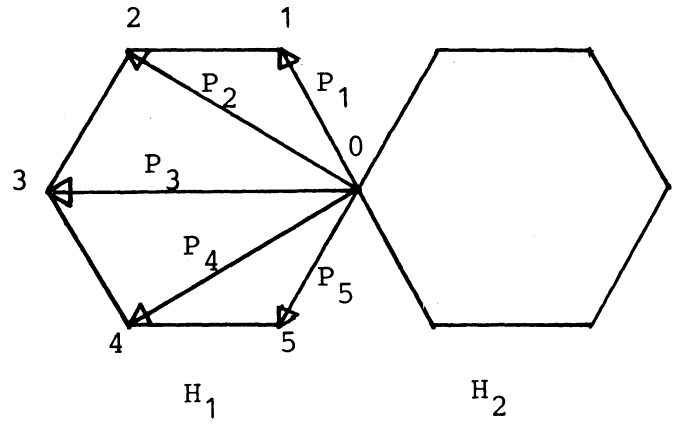


Figure 7

We first flip a fair coin to see which one of the hexagons H_1 and H_2 we want to go to next, and if H_1 comes up, we go to state i with probability p_i. I want these basic probabilities $p_1,..,p_5$ to be the same for all hexagons, and I also want them to be invariant under the obvious symmetries; hence $p_1=p_5$ and $p_2=p_4$. For technical reasons that I will return to in a moment, I also want them to decrease with distance; i.e. $p_3<p_2<p_1$. Hence to specify a random walk on S_n, all I need is to choose a triple (p_1,p_2,p_3) from the set

$$\mathcal{P} = \{(p_1,p_2,p_3) : p_1>p_2>p_3>0 \text{ and } 2p_1+2p_2+p_3=1\}$$

The problem is that if I just choose any element from $\mathcal{P}$ and try to trace the paths of the resulting process on S_n, I shall soon find myself in all kinds of combinatorial difficulties. And even if I could solve these difficulties for the relatively simple example I'm studying here, it's hard to believe that the solution would generalize in a natural way to the axiomatic setting I'm really interested in. Thus instead of working with general elements from $\mathcal{P}$, I'll choose very carefully one that minimizes the combinatorial problems.

Here's how I'll do it: In Figure 8 we're back to S_1. Notice that I have labelled the vertices of (what used to be) S_0 in exactly the same way as I labelled the vertices in Figure 7. We now choose transition probabilities $(p_1,p_2,p_3) \in \mathcal{P}$, start the corresponding random walk on S_1 in 0, and let it run until it hits one of states 1, 2, 3, 4 or 5, then stop.

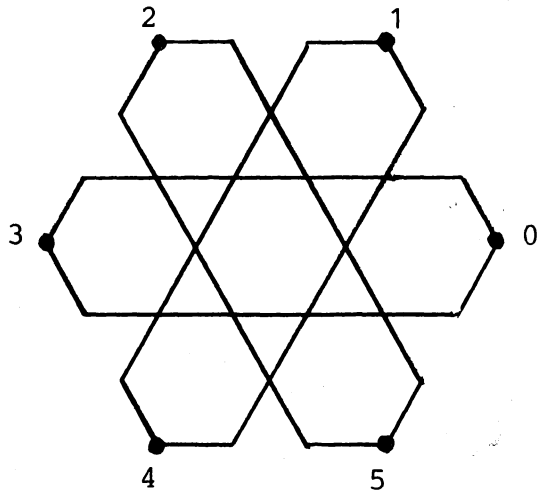


Figure 8

Let $\tilde{p}_i$ be the probability that it stops at site i. As we shall see in Chapter V, the map $(p_1,p_2,p_3) \to (\tilde{p}_1,\tilde{p}_2,\tilde{p}_3)$ is continuous and sends $\mathcal{P}$ into $\mathcal{P}$. Thus by the Lefschetz Fixed Point Theorem we can always choose the transition probabilities to be a fixed point for this map.

To see the advantage of such a choice of transition probabilities, consider the induced random walks B_n and B_N on S_n and S_N, respectively, for $N>n$. If we only observe how B_N hits points in S_n, the fixed point property ensures us that the two processes are indistinguishable; the same sequence of points is hit by the two processes with the same probability. Notice that in this argument I'm tacitly using an important property of the snowflake fractal; namely, that if B_N starts inside a hexagon in S_n, it can only leave that hexagon by passing through one of its vertices. An abstract formulation of this property will be one of the main axioms for "nested fractals" in the next section. Since B_N is just a refinement of B_n to a more finegrained lattice, it's fairly obvious that the sequence $\{B_n\}$ must have a limit as n goes to infinity. But to prove this, we must first understand

how to rescale time in the following sense: Assume that B_n typically uses time t_n to traverse one of the hexagons in S_n, and that B_N uses t_N to traverse a hexagon in S_N. If the two processes are to be indistinguishable when we also take time into account, we need to know the proper ratio between t_N and t_n; surely, t_N must be smaller than t_n, but by how much? I'll return to this question in a moment, but first I want to comment briefly on what has already been done.

Since I used the Lefschetz Fixed Point Theorem to choose my transition probabilities, a natural question is whether the map $(p_1, p_2, p_3) \to (\tilde{p}_1, \tilde{p}_2, \tilde{p}_3)$ has a unique fixed point. This question is all the more important as it's clear that different fixed points must necessarily lead to different processes. Unfortunately, I don't know the answer. All I can say is that if I had not introduced the condition $p_1 > p_2 > p_3$, there would clearly had been more than one fixed point; in addition to the one(s) found by applying Lefschetz's theorem to the domain $\mathcal{P}$, we would also have had the fixed point $p_1 = p_2 = 0$, $p_3 = 1$. The process induced by the last point is clearly degenerate in the sense that it doesn't span the whole fractal but only runs along diagonals, and the motivation for introducing the condition $p_1 > p_2 > p_3$ is to rule out degenerate processes of this kind. But still the possibility remains that some fractals may have more than one natural Brownian motion induced by different fixed points.

It's time to turn to the time scaling problem. I'll approach it in a slightly unconventional manner by introducing what I'll call <u>transition times</u>. Let us go back to Figure 7, assume that we are in state 0 and has decided to move to state i. The transition from 0 to i will take a certain amount of time

which I shall assume is governed by a random variable T_i with expectation t_i. As usual, I shall assume that the T_i's are invariant under the obvious symmetries; e.g. will T_1 and T_5 have the same distribution. I shall also assume that as I move along a path, the T_i's of the different steps will be independent. Hence all I need to specify are the distributions of the three <u>transition times</u> T_1, T_2, T_3.

Having chosen transition times T_1, T_2, T_3, I go back to Figure 8. As before, I start a random walk in state 0 and let it run until it hits one of the states 1,2,3,4 or 5, the only difference being that I now let each step take the time assigned to it by my chosen transition times. For each i=1,2,3,4 or 5, let $\tilde{T}_i$ be the distribution of the lifetime of the random walk given that it ends up in state i. Hence if $\tilde{t}_i$ is the expectation of $\tilde{T}_i$, $\tilde{t}_i$ is just the average time a particle which passes from 0 to i without hitting any of the other sites 1,2,3,4 or 5, spends on this journey.

Now the philosophy is the same as with the transition probabilities; I want to choose the T_i's such that the relationship between the T_i's and the $\tilde{T}_i$'s becomes as simple as possible. It's clearly absurd to hope for $\tilde{T}_i=T_i$, put perhaps it's possible to make $\tilde{T}_i$ and λT_i have the same distribution for some $\lambda>1$? This is indeed the case, and it's not even very hard to prove. The idea is as follows:

First observe that the expectations $\tilde{t}_i$ of the $\tilde{T}_i$'s only depend on the expectations t_i of the T_i's. In fact, the map

$(t_1, t_2, t_3) \to (\tilde{t}_1, \tilde{t}_2, \tilde{t}_3)$ is linear and the corresponding matrix has only positive elements. By Perron-Frobenius theory there is a largest, positive eigenvalue λ corresponding to an eigenvector $(\bar{t}_1, \bar{t}_2, \bar{t}_3)$ with positive components. If we give the space $\mathcal{T}$ of all distributions (T_1, T_2, T_3) with expectation $\bar{t}_1, \bar{t}_2, \bar{t}_3$ a suitable norm, the map sending (T_1, T_2, T_3) to $(\lambda^{-1} \tilde{T}_1, \lambda^{-1} \tilde{T}_2, \lambda^{-1} \tilde{T}_3)$ becomes a contraction, and hence has a unique fixed point $(\bar{T}_1, \bar{T}_2, \bar{T}_3)$.

Of course, the eigenvector $(\bar{t}_1, \bar{t}_2, \bar{t}_3)$ and the corresponding fixed point $(\bar{T}_1, \bar{T}_2, \bar{T}_3)$ is only determined up to a positive factor, and to fix the scale once and for all, I'll assume that $2p_1\bar{t}_1+2p_2\bar{t}_2+p_3\bar{t}_3=1$, i.e. the average time it takes to go from one point in S_0 to another one is always one.

The eigenvalue λ will be called the <u>time scaling factor</u>; it's one of the important parameters of the theory on a par with the volume scaling factor μ and the linear scaling factor ν introduced earlier.

I can now define a modified version $\hat{B}_n : \Omega \times \mathbb{R}_+ \to S_n$ of the random walk B_n on S_n as follows: We begin by choosing fixed points $(\bar{p}_1, \bar{p}_2, \bar{p}_3)$ and $(\bar{T}_1, \bar{T}_2, \bar{T}_3)$ as above. Once this has been done, the process is best explained informally by going back to Figure 7. Assume that the process has just arrived in state 0. To see where to go next, first flip a coin to decide which of the hexagons H_1 or H_2 to choose (if 0 only belongs to one hexagon, we can obviously skip this step). If H_1 is chosen, pick a state

$i=1,2,3,4$ or 5 randomly according to the distribution $(\bar{p}_1, \bar{p}_2, \bar{p}_3)$ (recall that by symmetry $\bar{p}_4=\bar{p}_2$ and $\bar{p}_5=\bar{p}_1$). Having decided which state i to go to, we must now wait for a random time before we are actually allowed to move there. The waiting time is distributed as $\lambda^{-n}\bar{T}_i$, and once it's over, we jump immediately to state i, and then repeat the procedure.

Let us compare $\hat{B}_n$ to $\hat{B}_N$ for $N>n$. As long as we only consider how these processes hit points in S_n, they are clearly indistinguishable not only with respect to which points are hit and in which order, but also with respect to the time when the hitting takes place. Thus it is intuitively obvious that the sequence $\{\hat{B}_n\}$ must have a limit B as n tends to infinity, and this can be established formally in a number of different ways (the theory of weak convergence certainly comes to mind). I prefer to use nonstandard analysis and simply define B to be the standard part of $\hat{B}_N$ for an infinitely large, nonstandard integer N.

As I'll prove later, B is a strong Markov process with continuous paths. (Note, by the way, that the approximants $\hat{B}_n$ are <u>not</u> usually Markov as information about how long the process has been waiting at a site, may give some indication of where it's likely to go next. As n goes to infinity, the waiting times get shorter, and $\hat{B}_n$ becomes "more and more Markov"). Another property that is intuitively clear from definition is that the process

$$(\omega,t) \to \nu B(\omega, \frac{t}{\lambda})$$

is essentially a copy of B; "essentially" meaning that we have to be a little careful where we start the process and what we do with

it when it hits the boundaries of the fractal. This scaling property should be compared to the well-known scaling property of ordinary Brownian motion b which says that

$$(\omega,t) \to \lambda^{\frac{1}{2}} b(\omega, \frac{t}{\lambda})$$

is a copy of b. If we write ν as $\lambda^{\log\nu/\log\lambda}$, we see that the difference between the two laws is that the exponent $\frac{1}{2}$ in the classical case has been replaced by $\log\nu/\log\lambda$ in the fractal case. Another way of considering the same phenomenon is by observing that the random walk $\hat{B}_n$ on S_n has space-increments of order of magnitude $\Delta x=\nu^{-n}$ and time-increments of size $\Delta t=\lambda^{-n}$. Hence $\Delta x=\Delta t^{\log\nu/\log\lambda}$ replaces the classical relation $\Delta x=\Delta t^{\frac{1}{2}}$. As I mentioned in the Introduction, a third place where this phenomenon occurs is in connection with Weyl's formula

$$n(\alpha) \sim \alpha^{d\cdot\frac{1}{2}} \qquad \text{as } \alpha \to \infty,$$

for the asymptotic distribution of eigenvalues of the Laplacian; also in this case the classical $\frac{1}{2}$ is replaced by the fractal $\log\nu/\log\lambda$ making the formula come out as

$$n(\alpha) \sim \alpha^{d\cdot\log\nu/\log\lambda} \qquad \text{as } \alpha \to \infty$$

(throughout the paper I shall write $f(t) \sim g(t)$ to denote that f and g diverge at the same rate; i.e., there are constants C and K such that $Cf(t) \leq g(t) \leq Kf(t)$ for all t).

Let me try to give a heuristic argument for this result (it's basically the same argument which in Chapter IX will be expanded into a full proof). For each $N\in\mathbb{N}$ and $x\in S_N$, let $q_N^t(x,x)$ be the probability that $\hat{B}_N(t)=x$ given that $\hat{B}_N(0)=x$, and define a trace-function Tr_N by

$$Tr_N(t)= \sum_{x\in S_N} q_N^t(x,x).$$

If $n<N$, S_N consists of μ^n scaled copies of S_{N-n}; call them $S^1_{N-n}, \ldots, S^{\mu^n}_{N-n}$. When $\hat{B}_N$ is restricted to one of these copies S^i_{N-n}, it looks exactly like $\hat{B}_{N-n}$ except that time has been rescaled by a factor λ^n and space by a factor ν^{-n}. If t is small, only a few particles will have time to pass from one copy S^i_{N-n} to another S^j_{N-n}, and hence

$$Tr_N(t) \approx \mu^n \, Tr_{N-n}(\lambda^n t).$$

Let $\{T^t\}$ be the Markov semi-group governing the Brownian motion on S. If we fix n and let N go to infinity, $Tr_N(t)$ and $Tr_{N-n}(t)$ both converge to $\text{trace}(T^t)$, and hence the formula above ought to turn into

$$\text{trace}(T^t) \approx \mu^n \, \text{trace}(T^{\lambda^n t}),$$

with better approximation the smaller t gets. The obvious solution to the functional equation

$$f(t) = \mu^n f(\lambda^n t) \qquad \text{for all } n \in \mathbb{N}$$

is $f(t) = t^{-\log\mu/\log\lambda}$, and it's easy to see that all other (nontrivial) solutions diverge at the same rate when $t \downarrow 0$. Hence $\text{trace}(T^t)$ ought to diverge as $t^{-\log\mu/\log\lambda}$ when $t \downarrow 0$. By Tauberian theory, this means that

$$n(\alpha) \sim \alpha^{\log\mu/\log\lambda} = \alpha^{d \cdot \log\nu/\log\lambda}, \qquad \text{as } \alpha \to \infty$$

(recall that $d = \log\mu/\log\nu$).

Even for a heuristic explanation, this argument may seem rather dubious, but by replacing the "approximate equal" sign $\approx$ by the appropriate inequalities, it will be possible to turn it into a rigorous proof.

A natural question at this point is whether it could actually be the case that $\log\nu/\log\lambda$ is always equal to $1/2$; i.e. $\lambda=\nu^2$. The answer is "no" as simple calculations show; e.g., $\nu=2$ and $\lambda=5$ for the Sierpinski gasket. In fact, the normal situation is that $\log\nu/\log\lambda$ is strictly less than $1/2$.

I'll end my informal description of Brownian motion on the snowflake fractal S here. As I have already said, the purpose of this chapter has been to give the reader an intuitive understanding of the quite simple, basic ideas of the paper. When I now dive into the more technical aspects of the general theory, my hope is that this intuitive understanding will make it possible to maintain a sense of purpose and direction. I'll do my best to fulfil this hope by referring back to this chapter whenever I find it appropriate.

IV. Nested fractals.

Nested fractals form a subclass of the family of self-similar sets introduced by Hutchinson [18]. In this chapter, I'll first give a brief review of those parts of Hutchinson's work that are particularly relevant to this paper (in fact, all I shall need can also be found in Section 8.3 of Falconer's book [11]), and then I'll introduce the axioms for nested fractals and draw a few simple consequences.

Let r be a real number. A r-<u>similitude</u> in $\mathbb{R}^k$ is simply a map $\psi:\mathbb{R}^k \to \mathbb{R}^k$ of the form

$$\psi(x) = r^{-1}U(x) + a,$$

where $U: \mathbb{R}^k \to \mathbb{R}^k$ is a unitary, linear map and $a \in \mathbb{R}^k$ is a constant. Given a finite family $\Psi=\{\psi_1,..,\psi_\mu\}$ of r-similitudes and a set $A \subset \mathbb{R}^k$, define

$$\Psi(A) = \bigcup_{i=1}^{\mu} \psi_i(A),$$

and let $\Psi^n(A)=\Psi\circ\ldots\circ\Psi(A)$ be the result of applying this operation n times. Hutchinson proved that Ψ has a fixed point in the following sense:

<u>IV.1 Theorem</u>. Assume that $r>1$. Then there is a unique compact set $E \subset \mathbb{R}^k$ such that $E=\Psi(E)$. Moreover, if $A \subset \mathbb{R}^k$ is any compact set, the iterates $\Psi^n(A)$ converge to E in the Hausdorff metric.

To understand the relevance of this result, let us go back to the snowflake fractal in the previous section. If we let $\psi_1,..,\psi_7$ be affine maps sending the hexagon in Figure 5a) to each

one of the smaller hexagons in Figure 5b), it's easy to check that the only fixed point of the system $\Psi=\{\phi_1,..,\phi_7\}$ is the snowflake fractal S. Completely analogous remarks apply to the three examples in Chapter II. Hence systems of similitudes seem to be the perfect tools for studying the kind of fractals we are interested in.

IV.2 Definition. A self-similar fractal with volume scaling factor μ and linear scaling factor ν is a pair (Ψ,E) consisting of a system $\Psi=\{\phi_1,..,\phi_\mu\}$ of ν-similitudes and its unique fixed point E.

I shall usually abuse terminology by just referring to the "self-similar fractal E"; strictly speaking this is nonsense as the same fixed-point E can be obtained from different systems Ψ with different scaling factors, but whenever I use this turn of phrase, I shall assume that I have one particular system in mind.

In Chapter II, I announced that the Hausdorff dimension of a self-similar fractal would be $\log\mu/\log\nu$ provided the scaled copies $\phi_1(E),..,\phi_\mu(E)$ do not come too close to overlap. Here's a convenient technical formulation of this condition:

IV.3 Open set condition. A system $\Psi=\{\phi_1,..,\phi_\mu\}$ of ν-similitudes satisfies the open set condition if there is a non-empty, bounded, open set V such that

$$\Psi(V)=\bigcup_{i=1}^{\mu}\phi_i(V)\subset V,$$

with the union disjoint.

Of our examples in the two previous sections, the Sierpinski gasket, the nested cubes, and the snowflake fractal obviously satisfy the open set condition: in each case we can just take V

to be the interior of the original triangle, cube, or hexagon. The Koch-curve is slightly more complicated, but Figure 9b) shows that we can take V equal to the interior of the triangle in Figure 9a).

Figure 9

IV.4 Theorem. (Hutchinson [18]) If a self-similar fractal E with volume scaling factor μ and linear scaling factor ν is generated by a system Ψ satisfying the open set condition, then E has Hausdorff dimension

$$d = \frac{\log \mu}{\log \nu} .$$

This is the result I have been referring to repeatedly above and which justifies my estimations of the Hausdorff dimensions in Chapters II and III. Let me point out that in his paper, Hutchinson proves stronger versions both of this theorem and of Theorem IV.1 as he allows the similitudes $\psi_1, .., \psi_\mu$ to have different scaling factors $\nu_1, .. \nu_\mu$; in this generality, the dimension d is no longer explicitly given, but is determined as the unique solution to the equation

$$\sum_{i=1}^{\mu} \nu_i^{-d} = 1 .$$

I have not tried to incorporate the case of different scaling factors in this paper as I find the theory sufficiently complicated as it is, but such an extension is certainly a natural project for future research.

In the last section, I described Brownian motion on the snowflake fractal as the limit of random walks on certain graphs S_n as n tended to infinity. The graphs S_n were constructed by an iterative procedure where at each step a basic configuration S_0 was replaced by a finite number of scaled copies of itself. In the general theory, the rôle of S_0 will be played by a finite set F (corresponding to the vertices of the hexagon in Figure 5a)), and the S_n's will be obtained as $\Psi^n(F)$. According to the last part of Theorem IV.1, the iterates $\Psi^n(F)$ will always converge to E, and thus it may seem irrelevant which finite set F we choose to start with. On second thoughts, however, it becomes clear that it's advantageous to let each element in F be a fixed point for one of the maps $\psi_1,..,\psi_\mu$ as we shall then have $\Psi^m(F)\supset\Psi^n(F)$ when $m>n$, and

$$E = \overline{\bigcup_{m\in\mathbb{N}} \Psi^m(F)},$$

where the bar denotes topological closure. It turns out that not all fixed points are equally essential.

IV.5 Definition. A point $x\in\mathbb{R}^k$ is fixed point for the system $\Psi=\{\psi_1,..,\psi_\mu\}$ if $\psi_i(x)=x$ for at least one i. The set of all fixed points is denoted by F_0. An element $x\in F_0$ is an essential fixed point if there are two distinct numbers i and j, $1\leq i,j\leq\mu$, and an element $y\in F_0$ such that

$$\psi_i(x) = \psi_j(y).$$

The set of all essential fixed points is denoted by F.

Note that since $\psi_1,..,\psi_\mu$ are contractions, the system Ψ has μ (not necessarily distinct) fixed points.

Essential fixed points are important because they tell us how the different parts of the fractal are put together; inessential fixed points serve no such purpose. To see this difference more clearly, let us go back for a moment to the snowflake fractal in the last chapter. This fractal can be obtained from a system $\Psi=\{\psi_1,..,\psi_7\}$ where each ψ_i consists of a scaling by the factor 1/3 followed by a translation. The fixed points of this system are the six vertices of the original hexagon plus its center (recall Figure 5). The vertices are essential fixed points as they are mapped to points where second generation hexagons meet, while the center is an inessential fixed point as all its images lie safely couched inside smaller hexagons. As we saw in the last chapter, the vertices played a crucial part in the description of the random walk on S_n, but the center was never mentioned. An even more striking illustration of the need to distinguish between essential and inessential fixed points is provided by the two essential and the two inessential fixed points of the Koch-curve; the readers are invited to work out this example for themselves.

Assume now that $\Psi=\{\psi_1,..,\psi_\mu\}$ is a system of ν-similitudes ($\nu>1$); let E be the induced self-similar fractal, and let F be the set of essential fixed points. For any set $A\subset \mathbb{R}^k$, I shall write $A^{(n)}$ for $\Psi^n(A)$, and if $i_1,..,i_n$ is a sequence of integers between 1 and μ, define

$$A_{i_1,..,i_n} = \psi_{i_1}\circ \cdots \circ \psi_{i_n}(A)$$

An element in $\mathbb{R}^k$ is an n-<u>point</u> if it belongs to $F^{(n)}$, and a set of the form $F_{i_1,..,i_n}$ is called an n-<u>cell</u>. The n-<u>complex</u> <u>associated with the</u> n-<u>cell</u> $F_{i_1,..,i_n}$ is $E_{i_1,..,i_n}$

(I apologize for the unimaginative terminology and hope that associations from algebraic topology will not create problems).

In the snowflake example, an n-point is a vertex in S_n; an n-cell is the set of vertices of one of the minimal hexagons making up S_n; and the associated n-complex is just the part of the fractal that lies inside this hexagon. Note, however, that in general the relationship between an n-cell and the associated n-complex is more complicated than the simple containment relation suggested by the last phrase (take a look at the Koch-curve for an example); the best that can be said in untechnical language is, perhaps, that the n-complex is the part of the fractal that "descends" from the n-cell.

We have now reached the stage where we can begin to take a look at the axioms for nested fractals. The first axiom rules out disconnected structures such as the Cantor-set.

Axiom 1. Connectivity: Any two 1-cells C and C' are connected by a sequence $C_1,..,C_n$ of 1-cells in the sense that $C=C_1$, $C'=C_n$, and C_i and C_{i+1} have a point in common for all i.

In the discussion of the snowflake fractal in the last chapter, the symmetry of the construction played an essential role. The second axiom will guarantee that nested fractals have the required symmetry, but before I state it, it is convenient to introduce some notation. If x and y are points in $\mathbb{R}^k$, let

$$H_{x,y}=\{z\in\mathbb{R}^k: |z-x|=|z-y|\}$$

be the hyperplane midway between them, and let $U_{x,y}$ be reflection in $H_{x,y}$.

Axiom 2. Symmetry: If $x,y \in F$, then $U_{x,y}$ maps n-cells to n-cells. Any n-cell which contains elements on both sides of $H_{x,y}$ is mapped to itself.

When we studied the snowflake fractal in the last chapter, we made crucial use of the fact that a random walk on S_N could only leave a hexagon in $S_n (n<N)$ by passing through one of its vertices. Our third and last axiom will be an abstract formulation of this principle.

Axiom 3. Nesting: If $i_1,..,i_n$ and $j_1,..,j_n$ are distinct sequences of integers between 1 and μ, then

$$E_{i_1,..,i_n} \cap E_{j_1,..,j_n} = F_{i_1,..,i_n} \cap F_{j_1,..,j_n}.$$

In the physics literature, fractals satisfying a condition of this kind are often called "finitely ramified". The Nesting Axiom is the most important but also the most restrictive of our axioms; it excludes many fractals one feels should be part of a general theory. One example is the Sierpinski carpet:

IV.6 Example: The construction is shown in Figure 10; the square in part a) is replaced by the 8 white squares in part b) etc.

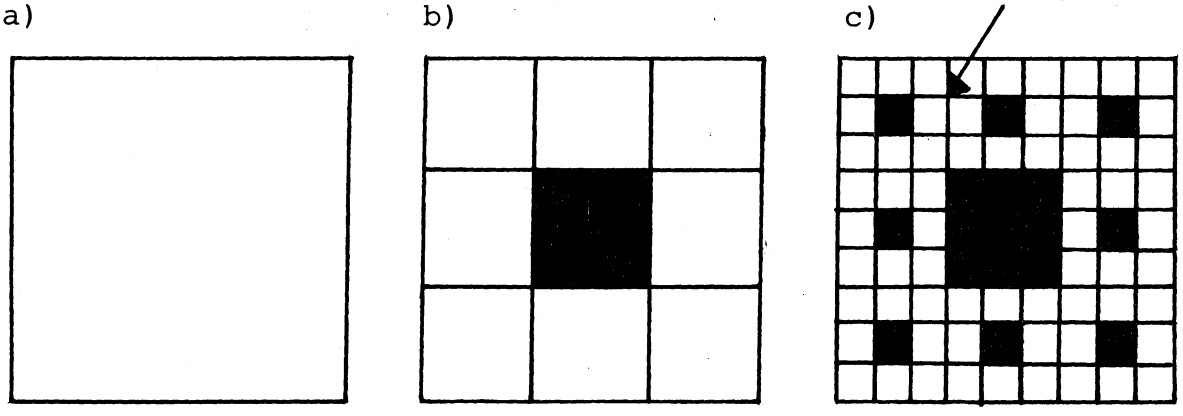


Figure 10

By pointing at a vertex which belongs to $E_1 \cap E_2$ but not to $F_1 \cap F_2$, the arrow in part c) shows that the Nesting Axiom is not satisfied. Nevertheless, Barlow and Bass [4] have constructed a Brownian motion on the Sierpinski carpet; to my knowledge, this is the only construction so far of a Brownian motion on a fractal which does not satisfy the axioms of this paper.

IV.7 Definition. A nested fractal is a self-similar fractal E derived from a system $\Psi=\{\psi_1,..,\psi_\mu\}$ of ν-similitudes which in addition to Axioms 1,2 and 3 satisfies the open set condition, and has at least two essential fixed points.

It is easy to check that all the examples in Chapter II and III are nested fractals in this sense.

The remainder of this section will be used to derive a few useful consequences from the axioms above. As often is the case with axiomatic theories, this will be a little pedantic and not always very exciting. The reader may want to skip the proofs at the first reading.

Let us begin by taking a look at a few simple consequences of the Symmetry Axiom. Define a symmetry to be a 1-similitude which maps n-cells to n-cells for all n. The first lemma tells us that we have all the symmetries we shall need.

IV.8 Lemma. Assume that $x,y,x'y' \in F$ and that $|x-y|=|x'-y'|$. Then there is a symmetry U such that $x=U(x')$ and $y=U(y')$.

Proof: Let $x''=U_{yy'}(x')$. If $x''=x$, we can simply let $U=U_{yy'}$.

If $x'' \neq x$, note that $|x-y|=|x''-y|$, and hence $U_{xx''}$ leaves y invariant. Thus $U=U_{xx''} \circ U_{yy'}$ does the job.

If we let

$$\delta_0=\min\{|x-y|: x,y\in F, x\neq y\},$$

then the shortest distance between two elements in the same n-cell is $\delta_n=\nu^{-n}\delta_0$. Two distinct points in E are n-<u>neighbors</u> if they are n-points belonging to the same n-cell; they are <u>nearest</u> n-<u>neighbors</u> if in addition the distance between them is δ_n. An n-<u>walk</u> is a sequence of n-points $s_1,..,s_p$ such that s_i and s_{i+1} are n-neighbors for all $i<p$; it is a <u>strict</u> n-<u>walk</u> if two consecutive elements always are nearest n-neighbors. Note that any point x in F must have a nearest 0-neighbor; this is just because if $y,z\in F$ and $|y-z|=\delta_0$, then $U_{x,y}(z)\in F$ and $|x-U_{x,y}(z)|=\delta_0$. It follows that any n-point has a nearest n-neighbor. In fact, something much stronger holds:

<u>IV.9 Lemma</u>. Two elements $x,y\in F$ are always connected by a strict 0-walk.

Proof: Assume not, and choose x and y to be the closest elements that can not be connected by a strict 0-walk. Let z be one of x's nearest neighbors. Clearly, $|z-y|>|x-y|$; if not there would be a strict walk connecting z and y, and this would immediately extend to one connecting x and y. Let $x'=U_{zy}(x)$, then $|x'-y|=\delta_0$. Also, since $|z-y|>|x-y|$, it follows by simple geometry that $|x-x'|<|x-y|$ (see Figure 11). But then x and x' are connected by a strict 0-walk which extends to one connecting x and y.

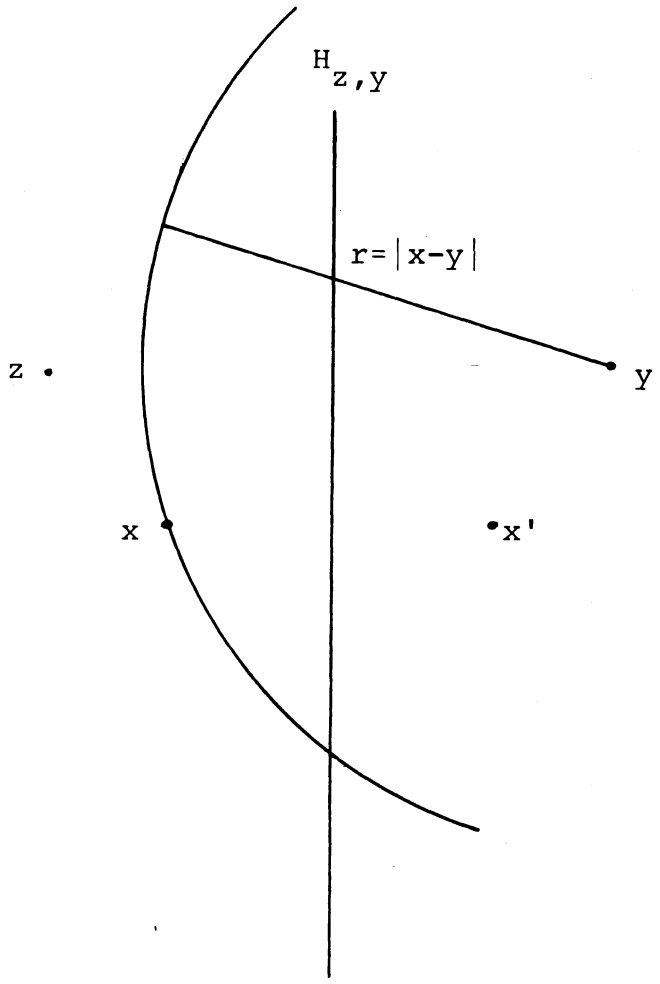


Figure 11

This result can be refined to:

IV.10 Lemma. If x,y and z are three distinct elements in F, there is a strict 0-walk which connects x and y and avoids z.

Proof: It's convenient to begin by introducing the following terminology. If x and y are elements of F, let the path-distance between x and y be the number of elements in the shortest strict 0-walk connecting x and y.

Assume now that the lemma is false, and among all triples x,y,z for which it fails, choose one for which the path-distance between x and y is as large as possible. Note that if $u \neq y$, then there must be a path which connects x and u and avoids y; if not, the path-distance between x and u would be strictly larger than the path-distance between x and y, and at the same time the lemma would fail for the triple x,u,y - hence contradicting the choice of x,y,z.

We are now ready for the main argument. Let $s_1,..,s_p$ be the shortest strict 0-walk connecting x and y. By assumption $z=s_k$ for (exactly) one $k<p$. For all i, let $s_i'=U_{y,z}(s_i)$; then $s_i',..,s_k'$ is a strict 0-walk connecting $x'=U_{y,z}(x)$ and y. Since $x'\neq y$, there is by the argument above a strict 0-walk $t_1,..,t_r$ which connects x and x' and avoids y. If $t_1,..,t_r$ does not contain z, then $t_1,..,t_{r-1}$, $s_1',..,s_k'$ is the walk we are looking for. On the other hand, if $t_j=z$ for some $j<r$, then t_r, $t_{r-1},..,t_j$ is a walk which connects x' and z and avoids y. Hence the image of this walk under $U_{y,z}$ is a strict 0-walk which connects x and y without hitting z.

Combining this lemma with the Connectivity Axiom, we get our first faintly interesting result:

<u>IV.11 Proposition</u>. Let $x,y\in F^{(1)}$. Then there is a strict 1-walk $s_1,..,s_p$ such that $x=s_1$, $y=s_p$ and $s_i\notin F$ for all i, $1<i<p$.

Before I give the proof, let me try to explain the significance of the result. If we return once again to the snowflake example in the last section, the proposition just says that starting in site 0 in Figure 8 and only moving along the edges of the hexagons, it is possible to get to any site 1,2,3,4 or 5 without passing through any other of these sites. Recall that when we assigned transition probabilities to the random walk on S_1, the probability p_1 of moving along the edge of a hexagon was necessarily strictly positive, and hence by the proposition all the composite transition probabilities $\tilde{p}_i$, $i=1,2,3$, must also be

strictly positive. Consequently, the associated process is nondegenerate in the sense that it runs all over the fractal. As we shall see in the next chapter, the proposition has exactly the same significance in the general case.

For the proof, I need a fact that I'll prove later in this chapter (Proposition IV.13): An element in F belongs to exactly one 1-cell. It follows from this that each 1-cell can contain at most one element from F (Corollary IV.14).

The proof begins by noting that according to the Connectivity Axiom there is a chain $C_1,..,C_n$ of 1-cells such that $x \in C_1$ and $y \in C_n$. For each $i<n$, let $t_i \in C_i \cap C_{i+1}$; by the result quoted in the last paragraph, t_i does not belong to F. Since C_i is just a scaled copy of F, Lemma IV.10 tells us that there is a strict 1-walk which connects t_i and t_{i+1} and which avoids the one point F and C_i may have in common. Sewing all these walks together, we get a strict 1-walk which connects t_1 and t_{n-1} and avoids F. Adding strict 1-walks connecting x and t_1 at the one end and t_{n-1} and y at the other end, we prove the proposition.

Although it is fairly obvious, let me state here for later reference the following reformulation of the Nesting Axiom in terms of walks.

IV.12 Proposition. Let n and N be positive integers, $n<N$. Assume that $s_1,s_2,..,s_p$ is an N-walk and that s_1 belongs to the n-complex $\hat{C}$ associated with an n-cell C. If s_k is the first element in the walk that does not belong to $\hat{C}$, then $s_{k-1} \in C$.

Proof: Assume that s_k belongs to the n-complex $\hat{D}$ associated with an n-cell D. Since the walk moves from one N-neighbor to another, s_{k-1} must also be an element of $\hat{D}$. But then

$$s_{k-1} \in \hat{C} \cap \hat{D} = C \cap D$$

by the Nesting Axiom.

The last topic I shall address in this section is the question of how many n-cells an element in E may belong to. I first treat the case needed for the proof of Proposition IV.11.

IV.13 Proposition. An element in F belongs to exactly one n-cell for each n.

Proof: Any $x \in F$ must belong to at least one n-cell C; just let $C = F_{i,..,i}$, where $\psi_i \in \Psi$ has x as a fixed point. If x also belongs to another n-cell $D = F_{j_1,...,j_n}$, then $x = \psi_{j_1,..,j_n}(y)$ for some $y \in F$. By Symmetry, y also belongs to two n-cells $F_{k,..,k}$ and $F_{l_1,..,l_n}$. But then x belongs to four 2n-cells $F_{i,..,i}$, $F_{i,..,i,j_1,..,j_n}$, $F_{j_1,..,j_n,k,..,k}$ and $F_{j_1,..,j_n,l_1,..,l_n}$.

Repeating the argument, we can make x an element of as many N-cells as we wish by just choosing N large enough. The idea of the proof is to show that this contradicts the open set condition IV.3.

If V is the open set in the open set condition, it follows from the last part of Theorem IV.1 that $E \subset \bar{V}$ (as usual, the bar denotes topological closure). For each $z \in F$, let B_z be the closed ball with center z and radius 1. Since $z \in \bar{V}$, the

intersection $B_z \cap V$ must have positive volume, and hence

$$\frac{\text{Vol } (B_z \cap V)}{\text{Vol } (B_z)} > \frac{1}{K}$$

for some positive integer K. Choose N so large that $x \in F$ belongs to K N-cells

$$\psi_{j_{1,1}} \circ \ldots\ldots \circ \psi_{j_{1,N}}(F), \ldots\ldots, \psi_{j_{K,1}} \circ \ldots\ldots \circ \psi_{j_{K,N}}(F),$$

and let z_i be the pre-image of x under the i-th of these maps. If B is the closed ball with center x and radius ν^{-N}, the family

$$\{\psi_{j_{i,1}} \circ \ldots\ldots \circ \psi_{j_{i,N}}(B_{z_i} \cap V)\}_{i \leq K}$$

is a disjoint collection of subsets of B each with a volume larger than $K^{-1} \times \text{vol}(B)$. As this is clearly impossible, the proposition is proved.

IV.14 Corollary. A 1-cell contains at most one element of F.

Proof: If a 1-cell $\psi_i(F)$ contains two elements of F, one of them - let us call it y - is not a fixed point for ψ_i. But then y is a fixed point for ψ_j for some $j \neq i$, and hence y belongs to two 1-cells $\psi_i(F)$ and $\psi_j(F)$. Since this contradicts the proposition, the corollary is proved.

The final results in this section generalize Proposition IV.13 to arbitrary elements in E. Let the multiplicity ρ of E be the largest number of 1-cells that meet in the same point. In all our standard examples the multiplicity is two, but for the modified Koch-curve in Figure 12 it is three.

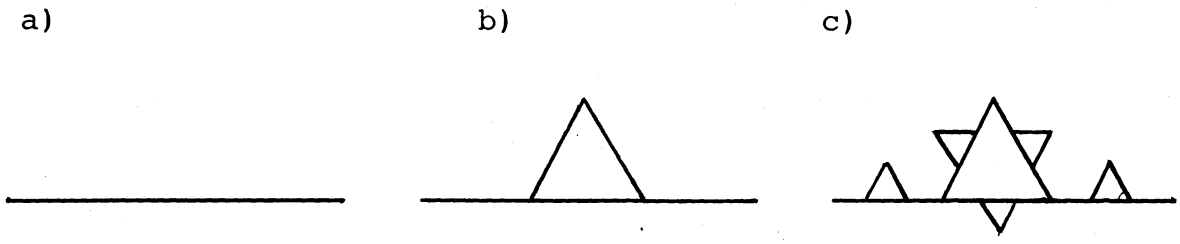


Figure 12

IV.15 Proposition. An element in E belongs to at most ρ n-complexes.

Proof: According to the Nesting Axiom, an element x in E can only belong to more than one n-complex if it is an n-point. I shall prove by induction on n that an n-point belongs to at most ρ n-comlexes. For $n=1$, this is just the definition of ρ (note that by the Nesting Axiom it doesn't matter whether we count cells or complexes). Assume that the assertion holds for $n<N$, and let x be an N-point belonging to the N-complexes

$$E_{i_{1,1},\ldots,i_{1,N}}, \ldots\ldots\ldots, E_{i_{m,1},\ldots,i_{m,N}}.$$

The proof now splits into two cases. Assume first that $i_{1,1}, i_{2,1}, \ldots, i_{m,1}$ are all equal to the same number i. Then $\psi_i^{-1}(x)$ is an (N-1)-point belonging to the (N-1)-complexes

$$E_{i_{1,2},\ldots,i_{1,N}}, \ldots\ldots\ldots, E_{i_{m,2},\ldots,i_{m,N}},$$

and thus $m<\rho$ by the induction hypothesis.

On the other hand, if there are numbers j and k such that $i_{j,1} \neq i_{k,1}$, then

$$(\divideontimes) \quad x \in E_{i_{j,1},\ldots i_{j,N}} \cap E_{i_{k,1},\ldots,i_{k,N}} \subset E_{i_{j,1}} \cap E_{i_{k,1}} = F_{i_{j,1}} \cap F_{i_{k,1}}$$

by the Nesting Axiom, and thus the point $y=\psi_{i_{j,1}}^{-1}(x)$ is an element of both F and the $(N-1)$-complex $E_{i_{j,2},\ldots,i_{j,N}}$. If there was another number $p\neq j$ such that $i_{p,1} = i_{j,1}$, then by the same argument y would also be an element of the $(N-1)$-complex $E_{i_{p,2},\ldots,i_{p,N}}$, and this is impossible by Proposition IV.13. Hence the numbers $i_{1,1},i_{2,1},\ldots i_{m,1}$ are all different, and by (⋇) above x belongs to all the 1-cells $F_{i_{1,1}},\ldots,F_{i_{m,1}}$. But then $m\leq\rho$ by the definition of ρ.

IV.16 Proposition. If x is an n-point, the number of N-complexes x belongs to is the same for all $N>n$.

Proof: Assume that

$$E_{i_{1,1},\ldots,i_{1,n}},\ldots\ldots\ldots,E_{i_{m,1},\ldots,i_{m,n}}$$

are the n-complexes x belongs to, and let

$$x=\psi_{i_{j,1}}\circ\ldots\ldots\circ\psi_{i_{j,n}}(x_j)$$

for each $j\leq m$. Since $x_j\in F$, it is the fixed point of one of the similitudes ψ_{i_j} in Ψ. Thus if $N>n$, x belongs to m N-complexes

$$E_{i_{j,1},\ldots\ldots i_{j,n},i_j\ldots i_j},\quad j\leq m.$$

Assume that x also belongs to an N-complex $E_{k_1,\ldots,k_N}$ that is

not on this list. Since

$$x \in E_{k_1,\ldots,k_N} \subset E_{k_1,\ldots,k_n},$$

the sequence $k_1,\ldots,k_n$ must be equal to $i_{j,1},\ldots i_{j,n}$ for some $j \leq m$. But then

$$x_j = (\psi_{i_{j,1}} \circ \ldots \circ \psi_{i_{j,n}})^{-1}(x) \in E_{k_{n+1},\ldots,k_N} \cap E_{j,..j},$$

and since $x_j \in F$, Proposition IV.13 tells us that $k_{n+1},\ldots,k_N$ equals $j,\ldots,j$.

The number of n-complexes an n-point x belongs to is called the <u>multiplicity</u> of x and is denoted by $\rho(x)$.

V. Transition probabilities

The basic transition probabilities $p_1,\ldots,p_5$ in Chapter III were supposed to be invariant under the natural symmetries of the hexagon S_0. In Lemma IV.8, I showed that if (x,y) and (x',y') are two pairs of elements from the set F of essential fixed points, then there is a symmetry mapping x' to x and y' to y if and only if $|x-y|=|x'-y'|$. Thus to study transition probabilities on nested fractals in general, it's natural to introduce an equivalence relation $\sim$ on the set

$$F^2-D=\{(x,y):x,y\in F \text{ and } x\neq y\}$$

by defining

$$(x,y)\sim(x',y') \text{ iff } |x-y|=|x'-y'|.$$

Let $c_1,\ldots,c_r$ be the equivalence classes of $\sim$ ordered by "increasing length" in the sense that if $(x,y)\in c_i$, $(u,v)\in c_j$, and $|x-y|<|u-v|$, then $i<j$. Pick an element $x\in F$. The <u>order</u> of the equivalence class c_i is the number

$$m_i=\text{card}\{y\in F:(x,y)\in c_i\}.$$

Note that by the Symmetry Axiom, m_i is independent of x. In the snowflake example in Chapter III, there are three equivalence classes c_1,c_2 and c_3, and their orders are $m_1=m_2=2$ and $m_3=1$. In analogy with this example, I define a set $\mathcal{P}$ of <u>basic transition probabilities</u> by

$$\mathcal{P}=\{(p_1,\ldots,p_r)\in R^r:p_1>p_2>\ldots>p_r>0 \text{ and } \sum_{i=1}^{r} m_i p_i=1\}$$

For each n, an element $(p_1,\ldots,p_r)\in\mathcal{P}$ induces a homogeneous Markov chain B_n on the set $F^{(n)}=\Psi^n(F)$ of n-points. I shall describe this process by specifying its transition probabilities $q_{x,y}$. If x and y are not n-neighbors (i.e. they either do not belong to the same n-cell, or x=y), then $q_{x,y}=0$. If x and y are n-neighbors, there exist a sequence $i_1,\ldots,i_n$ and two distinct points $x',y'\in F$ such that $x=\phi_{i_1}\circ\ldots\circ\phi_{i_n}(x')$ and $y=\phi_{i_1}\circ\ldots\circ\phi_{i_n}(y')$. The pair (x',y') belongs to an equivalence class c_i, and I define

$$q_{x,y}=\frac{p_i}{\rho(x)},$$

where $\rho(x)$ is the <u>multiplicity</u> of x; i.e. the number of n-cells x belongs to. Recall that by Proposition IV.16 the multiplicity is independent of n. A Markov chain B_n with transition probabilities $q_{x,y}$ is called a <u>Markov chain induced by</u> $(p_1,..,p_r)$. When I want to emphasize that the Markov chain starts at a point x, I shall denote it by B_n^x. The corresponding probability measure is then P_n^x.

As in Chapter III, I'm interested in the composite transition probabilities $\tilde{p}_1,\ldots,\tilde{p}_r$ induced by B_1. To define them, first introduce a stopping time τ_x for each $x\in F$ by

$$\tau_x(\omega)=\min\{i:\ B_1^x(i,\omega)\in F-\{x\}\}.$$

According to Proposition IV.11, it's possible to get from x to any other point in F, and thus by simple Markov chain theory τ_x is finite with probability one (see, e.g., Lemma V.3 below). If

$x,y \in F$, let $\tilde{p}_{x,y}$ be the probability that the process started at x is stopped at y, i.e.

$$\tilde{p}_{x,y} = P_1^x \{\omega: B_1^x(\tau_x(\omega),\omega)=y\}$$

V.1 Lemma. If $x,y,x',y' \in F$ and $|x-y|=|x'-y'|$, then $\tilde{p}_{x,y}=\tilde{p}_{x',y'}$.

Proof: By Lemma IV.8 there is a symmetry U such that $x=U(x')$ and $y=U(y')$. This map provides a one-to-one correspondence between paths of equal probability.

Given an equivalence class c_i of ~, choose a pair $(x,y) \in c_i$ and define

$$\tilde{p}_i=\tilde{p}_{x,y}.$$

By the lemma, $\tilde{p}_i$ is independent of the choice of (x,y). I shall call $\tilde{p}_1,..,\tilde{p}_r$ composite transition probabilities.

My aim in this section is to show that the map

$$\tilde{p}(p_1,\ldots,p_r)=(\tilde{p}_1,\ldots,\tilde{p}_r)$$

has a fixed point, and to do so I need to know that $\tilde{p}$ is continuous and maps $\mathcal{P}$ to $\mathcal{P}$.

V.2 Proposition. If $(p_1,\ldots,p_r) \in \mathcal{P}$, then $(\tilde{p}_1,\ldots,\tilde{p}_r) \in \mathcal{P}$.

Proof: For the purpose of this proof, it's convenient to modify the construction above slightly; instead of working with the Markov chain B_1, I shall use a chain Z with transition probabilities

$$\pi_{x,y}=\begin{cases} \tfrac{1}{2}q_{x,y} & \text{if } x \neq y \\ \tfrac{1}{2} & \text{if } x=y. \end{cases}$$

Thus Z behaves exactly like B_1 except that it occasionally hesitates for a while before it jumps from one state to the next. Note that if σ_x is the stopping time

$$\sigma_x(\omega)=\min\{i:Z(i,\omega)\in F-\{x\}\},$$

then

$$\tilde{p}_{x,y}=P^x\{\omega:Z(\sigma_x(\omega),\omega)=y\}.$$

To prove the proposition it suffices to show that if $x,y,y'\in F$ and $|x-y|<|x-y'|$, then $\tilde{p}_{x,y}>\tilde{p}_{x,y'}$. Let ζ be the set of paths Z can follow to get from x to y or y'. An element in ζ is a sequence $s_1,\ldots,s_n$ of 1-points satisfying:

(i) $s_1=x$ and s_n equals either y or y',

(ii) s_i and s_{i+1} belong to the same 1-cell for all $i<n$,

(iii) if $i<n$, then $s_i\notin F-\{x\}$

Let U be the symmetry $U_{y,y'}$ interchanging y and y'. If H is the closed half-space

$$H=\{z\in\mathbb{R}^k:|z-y|<|z-y'|\},$$

define $T:\mathbb{R}^k\to H$ by

$$T(u)=\begin{cases}u & \text{if } u\in H\\ U(u) & \text{if } u\notin H.\end{cases}$$

The <u>shadow</u> of a sequence $s_1,\ldots,s_n\in\zeta$ is the sequence $Ts_1,..,Ts_n$. Figure 13 shows the relationship between a sequence and its shadow. Note that the set

$$\zeta_0=\{Ts_1,\ldots,T_s:\ s_1,\ldots,s_n\in\zeta\}$$

of all shadows is a subset of ζ (this is the technical reason for working with Z and not B_1).

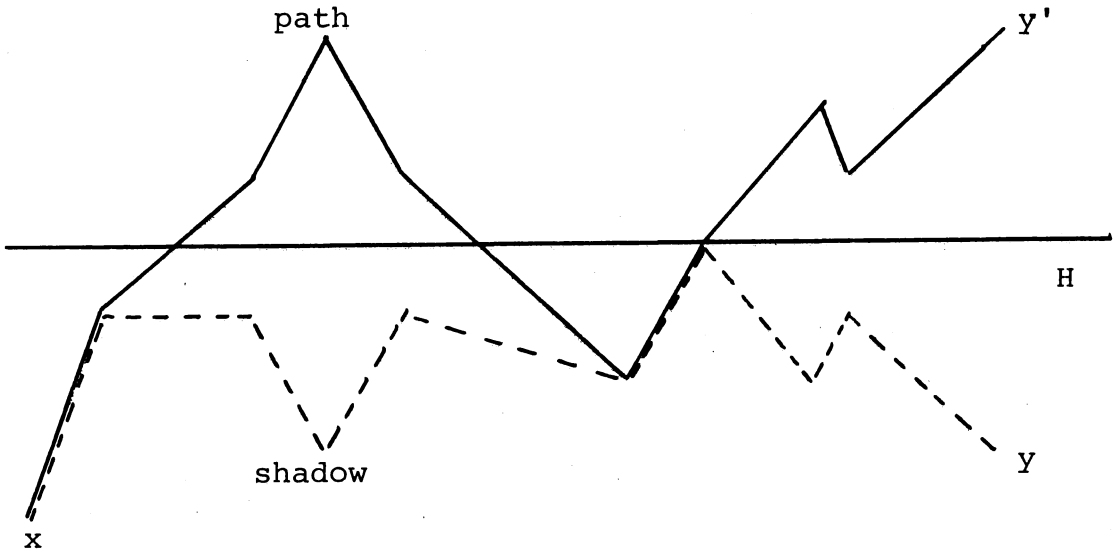


Figure 13

Two sequences are <u>equivalent</u> according to the equivalence relation $\equiv$ if they have the same shadow.

A sequence $s_1,\ldots,s_n$ is <u>leaving</u> H at i if $s_i \in H$ and $s_{i+1} \notin H$, and it is <u>entering</u> H at i if $s_i \notin H$ and $s_{i+1} \in H$. Clearly, $s_n = y$ if and only if $s_1,\ldots,s_n$ leaves and enters H the same number of times. Note that if $t_1,\ldots,t_n$ is the shadow of $s_1,\ldots,s_n$, then $s_1,\ldots,s_n$ can only leave or enter H at i if t_i and $U(t_{i+1})$ belong to the same n-cell; call such an index i <u>critical</u>. If i is critical, let $\hat{\pi}_{t_i,t_{i+1}} = \pi_{t_i,U(t_{i+1})}$, and observe that since $|t_i - U(t_{i+1})| \leq |t_i - t_{i+1}|$ and $(p_1,..,p_r) \in \mathcal{P}$, we must always have $\hat{\pi}_{t_i,t_{i+1}} \leq \pi_{t_i,t_{i+1}}$. Using Symmetry it's

easy to see that

$$\pi_{s_i,s_{i+1}} = \begin{cases} \hat{\pi}_{t_i,t_{i+1}} & \text{if } s_1,\dots,s_n \text{ leaves or enters H at } i \\ \pi_{t_i,t_{i+1}} & \text{otherwise.} \end{cases}$$

All that remains is the following simple calculation, where $\vec{t}=(t_1,t_2,\dots,t_n)$, $\vec{s}=(s_1,s_2,\dots,s_n)$ and $C(\vec{t})$ is the set of critical indices:

$$\tilde{p}_{x,y}-\tilde{p}_{x,y'}=$$

$$= \sum_{\vec{t}\in\zeta_0} \Bigg(\sum_{\substack{\vec{s}\equiv\vec{t} \\ s_n=y}} \pi_{s_1,s_2}\cdot\dots\cdot\pi_{s_{n-1},s_n} - \sum_{\substack{\vec{s}\equiv\vec{t} \\ s_n=y'}} \pi_{s_1,s_2}\cdot\dots\cdot\pi_{s_{n-1},s_n} \Bigg) =$$

$$= \sum_{\vec{t}\in\zeta_0} \prod_{i\notin C(\vec{t})} \pi_{t_i,t_{i+1}} \prod_{i\in C(\vec{t})} [\pi_{t_i,t_{i+1}} - \hat{\pi}_{t_i,t_{i+1}}] > 0,$$

where I've used the fact that $s_n=y'$ if and only if $\pi_{s_i,s_{i+1}} = \hat{\pi}_{t_i,t_{i+1}}$ for an odd number of indices i.

To prove that p is continuous, we shall need the following simple lemma.

V.3 Lemma. There exist constants C and θ, $0<\theta<1$, such that for all $n\in\mathbb{N}$ and all $x\in F$,

$$P_1^x \{\omega: \tau_x(\omega)>n\} < C\theta^n,$$

no matter which set $(p_1,\dots,p_r)\in\mathcal{P}$ of basic transition probabilities is used in the construction of B_1.

Proof: Let $k=(\sum_{i=1}^{r} m_i)^{-1}$. Note that since the sequence $p_1,\ldots,p_r$ is decreasing,

$$1=\sum_{i=1}^{r} m_i p_i \leq (\sum_{i=1}^{r} m_i)p_1,$$

and hence $p_1 \geq k$ whenever $(p_1,\ldots,p_r)\in\mathfrak{P}$. Proposition IV.11 tells us that if we just choose N large enough, any element u in $F^{(1)}$ is connected to $F-\{x\}$ by a strict 1-walk of length less than N, and thus if ρ is the multiplicity of E, we get

$$P_1^u\{\omega: \tau_u(\omega)\geq N\}\leq 1-(k/\rho)^N$$

By induction

$$P_1^u\{\omega: \tau_u(\omega)\geq mN\}\leq(1-(k/\rho)^N)^m$$

for all $m\in\mathbb{N}$ and from this the lemma follows easily.

IV.4 Corollary. The function $\tilde{p}(p_1,\ldots,p_r)=(\tilde{p}_1,\ldots,\tilde{p}_r)$ is continuous on $\mathfrak{P}$.

Proof: Given $x,y\in F$ and $n\in\mathbb{N}$, define

$$\tilde{p}^n_{x,y} = P_1^x\{\omega: B_1^x(\tau_x(\omega),\omega)=y \text{ and } \tau_x(\omega)\leq n\}.$$

For each $i\leq r$, let $\tilde{p}^n_i = \tilde{p}^n_{x,y}$ for $(x,y)\in c_i$. The functions $\tilde{p}^n(p_1,\ldots,p_r)=(\tilde{p}^n_1,\ldots,\tilde{p}^n_r)$ are clearly continuous; in fact, they are just polynomials of degree n. By the lemma, $\tilde{p}^n$ converges uniformly to $\tilde{p}$ in $\mathfrak{P}$, and hence $\tilde{p}$ is continous.

V.5 Theorem. The map $\tilde{p}:\mathfrak{P}\to\mathfrak{P}$ has a fixed point.

Proof: Since $\tilde{p}$ is continuous and maps the convex polyhedron $\mathfrak{P}$ to itself, this follows immediately from Lefschetz' or Brouwer's fixed point theorem (see [9]).

Note that if $(\bar{p}_1,\ldots,\bar{p}_r)$ is a fixed point for $\tilde{p}$, then all the components $\bar{p}_i$ are (strictly) positive; this follows from Proposition IV.11 and the fact that p_1 is always positive.

As I mentioned in Chapter III, the most important open question in the theory is whether the fixed point in Theorem V.5 is always unique. In the (very few) examples I have calculated there has never been more than one fixed point, but I really have no idea what the general situation is.

It should be clear from the exposition in Chapter III why the fixed points of $\tilde{p}$ are so important. To reformulate it in more technical terms, let x be an n-point and A a set of n-points. For each $N>n$, define

$$\sigma_A^N(\omega)=\min\{i:B_N^x(i,\omega)\in A\}$$

V.6 Proposition. If the basic transition probabilities $(\bar{p}_1,\ldots,\bar{p}_r)$ form a fixed point for $\tilde{p}$ and $A\subset F^{(n)}$, then

$$P_n^x\{B_n^x(\sigma_A^n)=y\}=P_N^x\{B_N^x(\sigma_A^N)=y\}$$

for all $y\in A$, $x\in F^{(n)}$, and $N>n$.

Proof: It suffices to prove the proposition for $N=n+1$, as we can then proceed by induction. Define a sequence $\{\tau_k\}$ of stopping times by letting $\tau_0=0$ and

$$\tau_{k+1}(\omega)=\min\{i>\tau_k(\omega):$$

$$B_{n+1}^x(i,\omega) \text{ is an n-point different from } B_{n+1}^x(\tau_k(\omega),\omega)\}$$

Define a new Markov chain Y by

$$Y(k,\omega)=B_{n+1}^x(\tau_k(\omega),\omega).$$

Combining Proposition IV.12 and the fact that $(\bar{p}_1,\ldots,\bar{p}_r)$ is a fixed point, it's easy to check that Y has the same distribution as B_n^x. Since A is a collection of n-points, this means that Y has the same hitting distribution as both B_n^x and B_{n+1}^x, and hence the proposition is proved.

VI. Transition times

So far our processes B_n are just Markov chains moving one step at each unit time. To construct a limit process, we must know how to rescale time as n grows large, and a convenient way of formulating this problem is in terms of random transition times as explained informally in Chapter III.

Let

$$\bar{\mathbb{R}}_+ = \{x \in \mathbb{R}: x>0\} \cup \{\infty\}$$

be the extended set of nonnegative real numbers. A set of <u>basic transition times</u> is just a sequence $(\tau_1,..,\tau_r)$ of completed Borel probability measures on $\bar{\mathbb{R}}_+$ - one for each equivalence class c_i. If $(x,y) \in c_i$, think of τ_i as the distribution of the time a random particle uses to get from x to y. I shall always assume that the τ_i's have finite second moments $\int x^2 d\tau_i(x)$, and I shall let t_i denote their expectations

$$t_i = \int x d\tau_i(x).$$

Fix sets $(p_1,\ldots,p_r) \in \mathfrak{P}$ and $(\tau_1,..,\tau_r)$ of basic transition probabilities and transition times, and for each n, let

$$B_n: \mathbb{N}_0 \times \Omega \to F^{(n)}$$

be a Markov chain induced by $(p_1,\ldots,p_r)$. I shall assume that $\{T_k\}$ is an increasing sequence of random variables satisfying certain conditions. In writing these conditions, it's convenient to let $(x,y) \in c_i^{(n)}$ denote the fact that x and y are n-points

of the form $x=\phi_{i_1}\circ\ldots\circ\phi_{i_n}(x')$ and $y=\phi_{i_1}\circ\ldots\circ\phi_{i_n}(y')$, where $x',y'\in F$ and $(x',y')\in c_i$.

VI.1 Conditions.

(i) $T_0=0$

(ii) The distribution of $T_{k+1}-T_k$ given that $B_n(k)=x$ and $B_n(k+1)=y$, where $(x,y)\in c_i^{(n)}$, is τ_i ; i.e. for all Borel sets A $P\{T_{k+1}-T_k\in A \mid B_n(k)=x \text{ and } B_n(k+1)=y\}=\tau_i(A)$.

(iii) The increments $T_{k+1}-T_k$, $k\in\mathbb{N}_0$, are independent of each other, and each of them is independent of $B_n(i)$ for $i\neq k,k+1$.

Define a process $\hat{B}_n: [0,\infty)\times\Omega\to F^{(n)}$ by

$$\hat{B}_n(t,\omega)=B_n(k,\omega),$$

where k is the largest integer such that $T_k(\omega)<t$. I shall refer to $\hat{B}_n$ as a random walk induced by $(p_1,\ldots,p_r)$ and $(\tau_1,..,\tau_r)$.

As explained in Chapter III, our aim is to find the "optimal choice" of the basic transition times $(\tau_1,..,\tau_r)$ by comparing them with composite transition times $(\tilde{\tau}_1,..,\tilde{\tau}_r)$. For each $x\in F$, let σ_x be the stopping time

$$\sigma_x(\omega)=\inf\{t\in\bar{\mathbb{R}}_+:\hat{B}_1^x(t)\in F-\{x\}\},$$

where $\hat{B}_1^x$ is just the random walk $\hat{B}_1$ started at x.

If y is another element of F, define $\tilde{\tau}_{x,y}$ to be the distribution of σ_x given that $\hat{B}_1^x(\sigma_x)=y$; i.e. $\tilde{\tau}_{x,y}$ is the

probability measure on $\bar{\mathbb{R}}_+$ given by

$$\tilde{\tau}_{x,y}(A)=P^x\{\sigma_x\in A \text{ and } \hat{B}_1^x(\sigma_x)=y\}/P^x\{\hat{B}_1^x(\sigma_x)=y\}$$

As usual, it is easy to see that $\tilde{\tau}_{x,y}=\tilde{\tau}_{x',y'}$ if $|x-y|=|x'-y'|$, and hence we may define $\tilde{\tau}_i$, $1\le i\le r$, by

$$\tilde{\tau}_i=\tilde{\tau}_{x,y}$$

whenever $(x,y)\in c_i$. These are the <u>composite transition times</u> referred to above.

<u>VI.2 Lemma</u>. The composite transition times $(\tilde{\tau}_1,..,\tilde{\tau}_r)$ have finite second moments.

Proof: It suffices to show that $E(\sigma_x^2)<\infty$. For each k, let $\Delta T_k=T_{k+1}-T_k$ and define

$$N_k(\omega)=\begin{cases}0 & \text{if } T_k(\omega)\ge\sigma_x(\omega)\\ 1 & \text{if } T_k(\omega)<\sigma_x(\omega).\end{cases}$$

Clearly, $\sigma_x(\omega)=\sum_{k=0}^{\infty} N_k(\omega)\Delta T_k(\omega)$, and thus

$$E(\sigma_x^2)=2\sum_{j<k}E(N_k(\omega)\Delta T_j(\omega)\Delta T_k(\omega))+\sum_{k=0}^{\infty}E(N_k(\omega)\Delta T_k(\omega)^2).$$

Let t and m be the largest of the expectations $t_i=\int x\,d\tau_i(x)$ and the second moments $m_i=\int x^2 d\tau_i(x)$ respectively. By Lemma V.3 there are constants C and θ, $0<\theta<1$, such that

$$P(N_k=1)<C\theta^k$$

for all k. Hence

$$E(\sigma_x^2)<2\sum_{j<k}C\theta^k t^2+\sum_{k=0}^{\infty}C\theta^k m = 2Ct^2\sum_{k=1}^{\infty}k\theta^k+Cm\sum_{k=0}^{\infty}\theta^k<\infty,$$

and the lemma is proved.

The expectations $\tilde{t}_1,..,\tilde{t}_r$ of the composite transition times $\tilde{\tau}_1,..,\tilde{\tau}_r$ can be computed fairly explicitly in terms of the expectations $t_1,..,t_r$ of the basic transition times $\tau_1,..,\tau_r$. To see how, fix an element $x\in F$, and for each $i<r$, choose $y_i\in F$ such that $(x,y_i)\in c_i$. Let Π_i be the set of all 1-walks $\vec{x}=x_0,\ldots,x_n$ such that $x_0=x, x_n=y_i$, and $x_j\notin F-\{x\}$ for all $j, 1<j<n$. The <u>length</u> $|\vec{x}|=n$ of the walk $\vec{x}$ is just the number of steps it has. Recall that q_{x_{j-1},x_j} is the probability that a particle which is in state x_{j-1} will jump to state x_j, and let t_{x_{j-1},x_j} be the average time it spends on this transition. Clearly, each t_{x_{j-1},x_j} will be one of the numbers $t_1,..,t_r$. Finally, let

$$\tilde{Q}_i=P_1^x\{\hat{B}_1^x(\sigma_x)=y_i\}.$$

We can now write

$$\tilde{t}_i=\tilde{Q}_i^{-1}\sum_{\vec{x}\in\Pi_i}\sum_{j=1}^{|\vec{x}|}t_{x_{j-1},x_j}\prod_{j=1}^{|\vec{x}|}q_{x_{j-1},x_j}, \tag{6.1}$$

which shows that the map

$$(\tilde{t}_1,..,\tilde{t}_r)=T(t_1,..,t_r)$$

is linear, and that the corresponding matrix has nonnegative entries. With just a little more thought it's easy to see that the entries must, in fact, be positive; this is equivalent to checking that for each pair (i,k), $1<i,k<r$, there is a path $x\in\Pi_i$ such that $(x_{j-1},x_j)\in c_k^{(1)}$ for at least one j, and in view of Corollary

IV.14 this is trivial; just choose x_1 such that $(x,x_1)\in c_k^{(1)}$, let $x_2=x$, and then follow any path from x to y.

VI.3 Proposition. There exist a positive number λ and a vector $(t_1,..,t_r)$ with positive components such that if $E(\tau_i)=t_i$ for each i, then $E(\tilde{\tau}_i)=\lambda t_i$ for each i. The constant λ is unique, and the vector $(t_1,...,t_r)$ is unique up to a scalar factor.

Proof: By the Perron-Frobenius Theorem (see, e.g., Seneta [31]), the linear map T has a largest eigenvalue λ which has multiplicity one and which corresponds to an eigenvector $(t_1,..,t_r)$ with positive components. The uniqueness follows from the Subinvariance Theorem of the Perron-Frobenius theory (Theorem 1.6 in Seneta [31]).

The constant λ in the proposition above is called the time scaling factor. As I mentioned in Chapter III, it is one of the crucial parameters of the theory on a par with the volume scaling factor μ and the linear scaling factor ν. It will follow from the general theory in Chapter IX that λ is the largest of these parameters (see Corollary IX.4); in fact $\lambda>\mu>\nu$.

The aim of this section is to show that it is possible to find distributions $(\tau_1,...,\tau_r)$ such that each $\tilde{\tau}_i$ is just the corresponding τ_i scaled by the time scaling factor λ;i.e.

$$\tilde{\tau}_i(A) = \tau_i\{x:\lambda x\in A\}$$

for all $i\le r$ and all Borel sets A. To prove this self-reproduction property is not very difficult - basically, it's a question of applying the fixed point theorem for contractions - but it will take me some work to set the stage properly.

If t is a positive real number, let $\mathcal{T}_t$ be the set of all completed Borel probability measures τ on $\bar{\mathbb{R}}_+$ with expectation t and finite second moment, and define $d_t: \mathcal{T}_t \times \mathcal{T}_t \to \mathbb{R}$ by:

$d_t(\sigma,\tau) = \inf\{(\int_\Omega |f-g|^2 dP)^{\frac{1}{2}} : (\Omega,P)$ is a probability space, and $f,g: \Omega \to \bar{\mathbb{R}}_+$ are random variables with distributions σ and τ, respectively$\}$

The only nontrivial part of checking that d_t is a metric on $\mathcal{T}_t$ is the triangle inequality

$$d_t(\sigma,\tau) \leq d_t(\sigma,\nu) + d_t(\nu,\tau).$$

To prove it, assume that $\varepsilon>0$ is given, and choose random variables $f,g_1: \Omega_1 \to \bar{\mathbb{R}}_+$, $g_2,h: \Omega_2 \to \bar{\mathbb{R}}_+$ such that f has distribution σ, g_1 and g_2 have distribution ν, h has distribution τ, and

$$\|f-g_1\|_{L^2(\Omega_1)} \leq d_t(\sigma,\nu)+\frac{\varepsilon}{2}, \quad \|g_2-h\|_{L^2(\Omega_2)} \leq d_t(\nu,\tau)+\frac{\varepsilon}{2}.$$

There then exist a probability space $(\tilde{\Omega},\tilde{P})$ and random variables $\tilde{f},\tilde{g},\tilde{h}: \tilde{\Omega} \to \bar{\mathbb{R}}_+$ such that the pair $(\tilde{f},\tilde{g})$ has the same joint distribution as (f,g_1), and $(\tilde{g},\tilde{h})$ has the same joint distribution as (g_2,h) (see, e.g., Stroyan and Bayod [33, Chapter IV])

But then

$$d_t(\sigma,\tau) \leq \|\tilde{f}-\tilde{h}\|_{L^2(\tilde{\Omega})} \leq \|\tilde{f}-\tilde{g}\|_{L^2(\tilde{\Omega})} + \|\tilde{g}-\tilde{h}\|_{L^2(\tilde{\Omega})} =$$

$$= \|f-g_1\|_{L^2(\Omega_1)} + \|g_2-h\|_{L^2(\Omega_2)} \leq d_t(\sigma,\nu)+d_t(\nu,\tau)+\varepsilon,$$

and since $\varepsilon>0$ is arbitrary, this proves the triangle inequality.

Nonstandard measure theory provides a straight-forward proof of the following crucial lemma. For the necessary background information, see Chapter 3 of [1].

VI.4 Lemma. $(\mathcal{T}_t, d_t)$ is complete.

Proof: If $\{\tau_n\}_{n\in\mathbb{N}}$ is a Cauchy-sequence, let $\{\hat{\tau}_n\}_{n\in{}^*\mathbb{N}}$ be its nonstandard version. Each $\hat{\tau}_n$ is an internal, $*$-countably additive measure on ${}^*\bar{\mathbb{R}}_+$. Choose an infinite element $N\in{}^*\mathbb{N}$, let $L(\hat{\tau}_N)$ be the Loeb-measure of $\hat{\tau}_N$, and define a measure τ on $\bar{\mathbb{R}}_+$ by

$$\tau(A)=L(\hat{\tau}_N)\ (st^{-1}(A)),$$

where st is the standard part map. It follows from general theory that τ is a complete Borel probability measure on $\bar{\mathbb{R}}_+$. The idea is that τ is the limit of $\{\tau_n\}$.

Let me first check that τ belongs to $\mathcal{T}_t$. Since $\{\tau_n\}$ is a Cauchy-sequence, the second moments $m_n=\int x^2 d\tau_n(x)$ are bounded by a constant $K\in\mathbb{R}$. By simple Loeb-measure theory

$$\int_{\bar{\mathbb{R}}_+} x^2 d\tau(x) = {}^*\int_{\bar{\mathbb{R}}_+} x^2 dL(\hat{\tau}_N)(x) < {}^*\int_{\bar{\mathbb{R}}_+} x^2 d\hat{\tau}_N(x) < K,$$

and hence τ has finite second moment. Moreover, since $\int x^2 d\hat{\tau}_N(x)$ is finite, it follows from Hölder's inequality that the function $x\mapsto x$ is S-integrable with respect to $\hat{\tau}_N$, and thus

$$\int_{\bar{\mathbb{R}}_+} x d\tau(x) = {}^*\int_{\bar{\mathbb{R}}_+} {}^\circ x dL(\hat{\tau}_N)(x) = {}^\circ{}^*\int_{\bar{\mathbb{R}}_+} x d\hat{\tau}_N(x) = {}^\circ t = t,$$

proving that $\tau\in\mathcal{T}_t$.

To prove that $\{\tau_n\}$ converges to τ, assume that $\varepsilon>0$ is given. Since $\{\tau_n\}$ is a Cauchy-sequence, there is a integer N_ε

such that $d_t(\tau_n, \tau_m) < \varepsilon$ for all $n,m > N_\varepsilon$. By definition of d_t, this means that there are random variables f_n^m, f_m^n: $\Omega \to \bar{\mathbb{R}}_+$ such that f_n^m has distribution τ_n, f_m^n has distribution τ_m, and $\|f_n^m - f_m^n\|_{L^2(\Omega)} < \varepsilon$. I can clearly assume that all the random variables f_n^m, $n,m \in \mathbb{N}$, are defined on the same sufficiently rich space Ω. Let $\{\hat{f}_n^m\}_{n,m \in {}^*\mathbb{N}}$ be the nonstandard version of the family $\{f_n^m\}_{n,m \in \mathbb{N}}$. If n is finite, the standard part $\hat{f}_n^m$ is a random variable on the Loeb-space $({}^*\Omega, L({}^*P))$ with distribution τ_n, and if N is the infinite number above, $\hat{f}_N^m$ is a random variable on $({}^*\Omega, L({}^*P))$ with distribution τ. Thus if $n > N_\varepsilon$,

$$d_t(\tau_n, \tau) < \left({}_*\int_\Omega (\hat{f}_n^N - \hat{f}_N^n)^2 dL({}^*P)\right)^{\frac{1}{2}} < \left({}_*\int_\Omega (\hat{f}_n^N - \hat{f}_N^n)^2 \, d^*P\right)^{\frac{1}{2}} < \varepsilon,$$

and the lemma is proved.

If $(t_1, \ldots, t_r)$ is the vector in Proposition VI.3 (-or, more correctly, one of the infinitely many vectors as $(t_1, .., t_r)$ is only determined up to a positive scalar factor), let $(\mathcal{T}, d)$ be the complete metric space defined by

$$\mathcal{T} = \mathcal{T}_{t_1} \times \ldots \times \mathcal{T}_{t_r}$$

and

$$d(\vec{\sigma}, \vec{\tau}) = \left(\sum_{i=1}^{r} d_{t_i}(\sigma_i, \tau_i)^4\right)^{1/4},$$

where $\vec{\sigma} = (\sigma_1, .., \sigma_r)$ and $\vec{\tau} = (\tau_1, .., \tau_r)$ are elements in $\mathcal{T}$. The choice of metric may seem a curious one, but d will later enter the argument in a perfectly natural way.

As before, let T be the linear map sending $(t_1,..,t_r)$ to $(\tilde{t}_1,...,\tilde{t}_r)$, and let $\Theta:\mathcal{T}\to\mathcal{T}$ be the map

$$\Theta(\tau_1,..,\tau_r) = (\Theta_1(\tau),...,\Theta_r(\tau)),$$

where each component $\Theta_i(\tau)$ is $\hat{\tau}_i$ scaled by λ^{-1}; i.e. for all Borel sets A

$$\Theta_i(\tau)(A)=\tilde{\tau}_i\{x:\lambda^{-1}x\in A\}.$$

To prove the main result of this section, it suffices to show that Θ has a unique fixed point:

VI.5 Theorem. There is a sequence $(\tau_1,..,\tau_r)$ of basic transition times such that the composite transition times $(\tilde{\tau}_1,..,\tilde{\tau}_r)$ are just $(\tau_1,..,\tau_r)$ scaled by the time scaling factor λ. The sequence $(\tau_1,..,\tau_r)$ is unique up to a scaling factor.

Proof: Fix an element $x\in F$, and for each $i\leq r$, choose $y_i\in F$ such that $(x,y_i)\in c_i$. Recall that Π_i is the set of all 1-walks $\vec{x}=x_0,...,x_n$ such that $x_0=x, x_n=y_i$ and $x_j\notin F-\{x\}$ for all $j, 1<j<n$. Thus if $B_1:\mathbb{N}_0\times\Omega\to F^{(1)}$ is the Markov chain induced by our chosen transition probabilities $(p_1,..,p_r)$, and τ_x is the stopping time

$$\tau_x(\omega)=\min\{i:B_1^x(i,\omega)\in F-\{x\}\},$$

then Π_i is the set of all paths the stopped process $B_1^x(n\wedge\tau_x(\omega),\omega)$ can follow from x to y_i. For each i, let

$$\Omega_i=\{\omega:B_1^x(\tau_x(\omega),\omega)=y_i\},$$

set $\tilde{Q}_i=P_1^x(\Omega_i)$, and define a probability measure Q_i on Ω_i by

$$Q_i(A)=\tilde{Q}_i^{-1}P_1^x(A).$$

I shall use E_i to denote expectation with respect to Q_i; thus E_i is the expectation given that $B_1(0)=x$ and $B_1(\tau_x)=y_i$. For each $\vec{x}\in\Pi_i$, define $E_{\vec{x}}$ to be the expectation given that $B_1(j)=x_j$ for all $j\leq|\vec{x}|$ in an analogous way.

Let $\{T_k\}$ and $\{T_k^*\}$ be two sequences of random variables satisfying Condition VI.1, but corresponding to two different sets $\vec{\tau}=(\tau_1,..,\tau_r)$, $\vec{\tau}^*=(\tau_1^*,..,\tau_r^*)$ of distributions in $\mathcal{T}$. Assume that $\{T_k\}$ and $\{T_k^*\}$ are chosen such that $\Delta T_k=T_{k+1}-T_k$ and $\Delta T_\ell^*=T_{\ell+1}^*-T_\ell^*$ are independent when $k\neq\ell$. By definition, $\Theta_i(\tau)$ is just the distribution of

$$\tilde{T}_i(\omega)=\lambda^{-1}T_{\tau_x}(\omega)$$

as a random variable on (Ω_i,Q_i), and, similarly, $\Theta_i(\tau^*)$ is the distribution of

$$\tilde{T}_i^*(\omega)=\lambda^{-1}T^*_{\tau_x}(\omega)$$

as a random variable on (Ω_i,Q_i). Observe that

(6.2) $E_i[(\tilde{T}_i-\tilde{T}_i^*)^2]=$

$$\lambda^{-2}\tilde{Q}_i^{-1}\sum_{\vec{x}\in\Pi_i}E_{\vec{x}}\left[(T_{|\vec{x}|}-T^*_{|\vec{x}|})^2\right]\prod_{j=1}^{|\vec{x}|}q_{x_{j-1},x_j}=$$

$$\lambda^{-2}\tilde{Q}_i^{-1}\sum_{\vec{x}\in\Pi_i}E_{\vec{x}}\left[\left(\sum_{j=1}^{|\vec{x}|}(\Delta T_{j-1}-\Delta T^*_{j-1})\right)^2\right]\prod_{j=1}^{|\vec{x}|}q_{x_{j-1},x_j}=$$

$$\lambda^{-2}\tilde{Q}_i^{-1}\sum_{\vec{x}\in\Pi_i}\sum_{j=1}^{|\vec{x}|}E_{\vec{x}}\left[(\Delta T_{j-1}-\Delta T^*_{j-1})^2\right]\prod_{j=1}^{|\vec{x}|}q_{x_{j-1},x_j},$$

where the last step uses that the random variables $\Delta T_{j-1} - \Delta T^*_{j-1}, j < |\vec{x}|$, are independent and have mean zero.

Given a path $\vec{x} = x, .., x_n$, the distribution of ΔT_{j-1} is just τ_k for some $k < r$; let us write τ_{x_{j-1}, x_j} for this τ_k. Similarly, the distribution of ΔT^*_{j-1} is $\tau^*_{x_{j-1}, x_j} = \tau^*_k$. By choosing the sequences $\{T_k\}$ and $\{T^*_k\}$ judiciously, we can get $E^{\rightarrow}_x [(\Delta T_{j-1} - \Delta T^*_{j-1})^2]$ as close to $d_{t_{x_{j-1}, x_j}}(\tau_{x_{j-1}, x_j}, \tau^*_{x_{j-1}, x_j})^2$ as we want to.

Thus the infimum of the right hand side of (6.2) over all sequences $\{T_k\}$, $\{T^*_k\}$ satisfying our conditions is

$$\lambda^{-2} \tilde{Q}^{-1}_i \sum_{\vec{x} \in \Pi_i} \sum_{j=1}^{|\vec{x}|} d_{t_{x_{j-1}, x_j}}(\tau_{x_{j-1}, x_j}, \tau^*_{x_{j-1}, x_j})^2 \prod_{j=1}^{|\vec{x}|} q_{x_{j-1}, x_j},$$

which according to (6.1) is nothing but the i-th coordinate of

$$\lambda^{-2} T(d_{t_1}(\tau_1, \tau^*_1)^2, \ldots, d_{t_r}(\tau_r, \tau^*_r)^2),$$

where T is the linear map sending $(t_1, \ldots, t_r)$ to $(\tilde{t}_1, .., \tilde{t}_r)$. At the same time, the left hand side of (6.2) is always at least as large as $d_{t_i}(\Theta_i(\tau), \Theta_i(\tau^*))^2$, and thus

$$(6.3)\quad (d_{t_1}(\Theta_1(\tau), \Theta_1(\tau^*))^2, \ldots\ldots, d_{t_r}(\Theta_r(\tau), \Theta_r(\tau^*))^2) <$$

$$\lambda^{-2} T(d_{t_1}(\tau_1, \tau^*_1)^2, \ldots\ldots, d_{t_r}(\tau_r, \tau^*_r)^2),$$

with the inequality holding componentwise. Since T's matrix has only positive entries, we can apply T to both sides of (6.3) and

keep the inequality. If we do this repeatedly, we get

$$(d_{t_1}(\theta_1^k(\tau),\theta_1^k(\tau^*))^2,\ldots\ldots,\ d_{t_r}(\theta_r^k(\tau),\theta_r^k(\tau^*))^2)$$

$$<\lambda^{-2k}T^k(d_{t_1}(\tau_1,\tau_1^*)^2,\ldots\ldots,\ d_{t_r}(\tau_r,\tau_r^*)^2)$$

for all natural numbers k. Since T is a Perron-Frobenius map with leading eigenvalue λ, T has an asymptotic expression

$$T^k=\lambda^k A+\mathcal{O}(k^s|\lambda_2|^k) \quad \text{as} \quad k\to\infty,$$

where A is a fixed matrix, s and λ_2 are fixed numbers, and $|\lambda_2|<\lambda$ (see Theorem 1.2 in Seneta [31]). Thus for some sufficiently large k, the linear map $\lambda^{-2k}T^k$ has operator norm C less than one, and hence

$$d(\theta^k(\vec{\tau}),\theta^k(\vec{\tau}^*))<C^{\frac{1}{2}}d(\vec{\tau},\vec{\tau}^*)$$

by (6.3) and the definition of the metric d. This means that θ^k is a contraction and has a unique fixed point $\vec{\tau}_0$. But if $\vec{\tau}_0$ is a fixed point for θ^k so is $\theta(\vec{\tau}_0)$, and since θ^k has only one fixed point, $\vec{\tau}_0=\theta(\vec{\tau}_0)$ and is thus a fixed point for θ. Since any other fixed point of θ automatically is a fixed point of θ^k, $\vec{\tau}_0$ is unique. The proof is complete.

Call a pair $(\vec{p},\vec{\tau})$ of transition probabilities $\vec{p}=(p_1,..,p_r)$ and transition times $\vec{\tau}=(\tau_1,..,\tau_r)$ <u>stable</u> if $\vec{p}$ is a fixed point for the map $\tilde{p}$ in Theorem V.5 and $\vec{\tau}$ satisfies Theorem VI.5.

For stable pairs $(\vec{p},\vec{\tau})$ it is possible to extend Proposition V.6 by also taking time into account. If $\hat{B}_n$ is the random walk $(\vec{p},\vec{\tau})$ induces on $F^{(n)}$, and A is a collection of n-points, let

$$\tau_A^n(\omega)=\inf\{t:\ \hat{B}_n(t,\omega)\in A\}$$

VI.6 Proposition. If $(\vec{p},\vec{\tau})$ is stable, $x\in F^{(n)}$, and A is a collection of n-points, then

$$P_n^x\{\hat{B}_n^x(\tau_A^n)=y\}=P_N^x\{\hat{B}_N^x(\tau_A^N)=y\}$$

for all $y\in A$ and all $N>n$. Moreover,

$$P_n^x\{\lambda^{-n}\tau_A^N(\omega)\in B\,|\,\hat{B}_n^x(\tau_A^n)=y\}$$

$$= P_N^x\{\lambda^{-N}\tau_A^N(\omega)\in B\,|\,\hat{B}_N^x(\tau_A^N)=y\}$$

for all Borel sets B.

Proof: The first part of the statement is just a reformulation of Proposition V.6. It clearly suffices to prove the second part for $N=n+1$, as we can then proceed by induction. Define a sequence $\{\tilde{T}_k\}$ of stopping times by $\tilde{T}_0=0$ and

$$\tilde{T}_{k+1}(\omega)=\inf\{t>\tilde{T}_k(\omega):\ \hat{B}_{n+1}^x(t,\omega)\in F^{(n)}-\{\hat{B}_{n+1}^x(\tilde{T}_k(\omega),\omega)\}\}.$$

By Theorem VI.5, $\{\lambda^{-1}\tilde{T}_k\}$ is a sequence of random variables satisfying Condition VI.1, and hence the process

$$\hat{C}_n(t,\omega)=\hat{B}_{n+1}(\tilde{T}_k(\omega),\omega)$$

where k is the largest integer such that $\lambda^{-1}\tilde{T}_k<t$, is a copy of $\hat{B}_n$ with the same hitting time distribution. The proposition now follows easily.

I'll close this section with two simple examples of how one can use the fact that stable transition times are fixed points for the map Θ to obtain information about their properties. If $x_0,x_1,..,x_q$ is the shortest 1-walk leading from one element $x=x_0$ in F to another $y=x_q\in F$, call q the connection factor of the fractal. Given a stable set $\vec{\tau}=(\tau_1,..,\tau_r)$ of transition times,

let $\{T_k\}$ be an associated sequence of random variables satisfying Condition VI.1, and define a function $\zeta:[0,\infty)\to R$ by

$$\zeta(t)=P\{T_1(\omega)<t\}$$

VI.7 Proposition. If $(\vec{p},\vec{\tau})$ is stable and $\beta=\log q/\log\lambda$, then there is a constant $K>1$ such that

$$\zeta(t)<K^{-t^{-\beta}} \quad \text{for all } t\in(0,1]. \tag{6.4}$$

Proof: It's easy to see that there must be a constant K such that (6.4) holds for all $t\in(\lambda^{-1},1]$. Define

$$f(t)=K^{-t^{-\beta}}$$

and note that

$$f(t)=f(\lambda t)^q \quad \text{for all } t>0.$$

If I can show that

$$\zeta(t)<\zeta(\lambda t)^q \quad \text{for all } t>0,$$

then the following simple argument will complete the proof: Given a $t\in(0,1]$, choose $n\in\mathbb{N}$ such that $\lambda^n t\in(\lambda^{-1},1]$ and note that

$$\zeta(t)<\zeta(\lambda^n t)^{q^n}<f(\lambda^n t)^{q^n}=f(t).$$

To prove that $\zeta(t)<\zeta(\lambda t)^q$, choose an element $x\in F$ and let

$$\sigma_x(\omega)=\inf\{t:\hat{B}_1^x(t,\omega)\in F-\{x\}\}.$$

Since $(\vec{p},\vec{\tau})$ is stable

$$\zeta(t)=P\{\omega:T_1(\omega)<t\}=P^x\{\omega:\sigma_x(\omega)<\lambda t\},$$

and by definition of q

$$P^x\{\omega:\sigma_x(\omega)<\lambda t\}<P\{\omega:T_q(\omega)<\lambda t\}$$

$$<\prod_{i=0}^{q-1} P\{\omega:\Delta T_i(\omega)<\lambda t\}=P\{\omega:T_1(\omega)<\lambda t\}^q=\zeta(\lambda t)^q.$$

Hence $\zeta(t)<\zeta(\lambda t)^q$, and the proof is complete.

Let us agree that a measure m on $[0,\infty)$ has a C^∞-density if there is an infinitely differentiable function $f:[0,\infty)\to\mathbb{R}$ such that

$$m(A)=\int_A f(x)dx$$

for all Borel sets A.

VI.8 Proposition. If $(\vec{p},\vec{\tau})$ is a stable pair, then each τ_i has a C^∞-density.

Proof: I'll use the following convention for Fourier transforms $\hat{f}$ and inverse Fourier transforms g:

$$\hat{f}(y)=\int_{-\infty}^{\infty} e^{-2\pi iyx} f(x)dx$$

$$g(x)=\int_{-\infty}^{\infty} e^{2\pi ixy} g(y)dy.$$

To compute the Fourier transform of the measure τ_i, choose two points $x,y\in F$ such that $(x,y)\in c_i$, and let

$$\sigma(\omega)=\inf\{t:\ \hat{B}_1(t,\omega)\in F-\{x\}\}.$$

Of course, τ_i is the distribution of $\lambda^{-1}\sigma$ given that $\hat{B}_1(0)=x$ and $\hat{B}_1(\sigma)=y$, and thus

$$\hat{\tau}_i(y)=\int^{\infty} e^{-2\pi iyx}\, d\tau_i(x)=$$

$$= E(\ e^{-2\pi iy\lambda^{-1}\sigma}\ |\ \hat{B}_1(0)=x \text{ and } \hat{B}_1(\sigma)=y).$$

$$= \tilde{Q}_i^{-1} \sum_{\vec{x}\in\Pi_i} E_{\vec{x}}\,(e^{-2\pi iy\lambda^{-1}\sigma}) \prod_{j=1}^{|\vec{x}|} q_{x_{j-1},x_j}$$

in the notation used earlier in this section. Let $\{\sigma_k\}_{k\in\mathbb{N}}$ be the sequence of stopping times defined by $\sigma_0\equiv 0$ and

$$\sigma_{k+1}(\omega)=\inf\{t>\sigma_k(\omega):\ \hat{B}_1(t,\omega)\neq\hat{B}_1(\sigma_k(\omega),\omega)\}.$$

Then

$$\hat{\tau}_i(y) = \tilde{Q}_i^{-1} \sum_{\vec{x}\in\Pi_i} E_{\vec{x}} \left(e^{-2\pi i \frac{y}{\lambda} \sum_{k=1}^{|\vec{x}|} (\sigma_k - \sigma_{k-1})}\right) \prod_{j=1}^{|\vec{x}|} q_{x_{j-1},x_j}$$

$$= \tilde{Q}_i^{-1} \sum_{\vec{x}\in\Pi_i} \prod_{k=1}^{|\vec{x}|} E_{\vec{x}} \left(e^{-2\pi i \frac{y}{\lambda} (\sigma_k - \sigma_{k-1})}\right) \prod_{j=1}^{|\vec{x}|} q_{x_{j-1},x_j}.$$

Given the path $\vec{x}$, each $\sigma_k - \sigma_{k-1}$ has distribution $\tau_{i_{x_k}}$ for some $i_{x_k} \in \{1,\dots,r\}$, and hence

$$(6.5)\quad \hat{\tau}_i(y) = \tilde{Q}_i^{-1} \sum_{\vec{x}\in\Pi_i} \prod_{k=1}^{|\vec{x}|} \hat{\tau}_{i_{x_k}}\left(\frac{y}{\lambda}\right) \prod_{j=1}^{|\vec{x}|} q_{x_{j-1},x_j}.$$

I now introduce the function

$$F(y) = \max_{1\le i\le r} |(\hat{\tau}_i(y)|.$$

Since each $\hat{\tau}_i$ is the Fourier transform of a probability measure, $F(y)\le 1$ for all y, and hence it follows from (6.5) that

$$F(y)\le F\left(\frac{y}{\lambda}\right)^q,$$

where q still is the connection factor. Choose $K>1$ so small that

$$F(y)\le K^{-|y|^{\log q/\log\lambda}} \quad \text{when} \quad |y|\in(\lambda^{-1},1]$$

and define

$$G(y) = K^{-|y|^{\log q/\log\lambda}}.$$

Note that

$$G(y) = G\left(\frac{y}{\lambda}\right)^q \quad \text{for all } y.$$

If $|y|>1$, there is an $n\in\mathbb{N}$ such that $|\lambda^{-n}y|\in(\lambda^{-1},1]$, and thus

$$F(y)\le F(\lambda^{-n}y)^{q^n} \le G(\lambda^{-n}y)^{q^n} = G(y).$$

To prove the proposition it only remains to observe that since $\hat{\tau}_i(y) \leq F(y) \leq K^{-|y|^{\log q/\log \lambda}}$ when $|y| > 1$, the function $y^n \hat{\tau}_i(y)$ is bounded and integrable for all positive powers n. By basic Fourier theory it follows that the inverse Fourier transform of $\hat{\tau}_i$ is a function with derivatives of all orders.

Remarks: In their paper on Brownian motion on the Sierpinski gasket [5], Barlow and Perkins use the theory of branching processes to find stable transition times. It would be interesting to know if their methods generalize to the present situation, and, if so, whether this could lead to sharper estimates. As we shall see in the next chapter, it would be particularly interesting to know the best value of β for which formula (6.4) holds.

VII. Brownian motion on nested fractals.

Just as ordinary random walks in $\mathbb{R}^k$ converge to Brownian motion, the random walks introduced in the previous chapter converge to continuous time processes on E. I shall refer to these limit processes as Brownian motions on E, and the purpose of the present chapter is to construct them and study them in some detail.

Throughout the chapter E will be a nested fractal induced by a system $\Psi=\{\psi_1,..,\psi_\mu\}$ of ν-similitudes, and $(\vec{p},\vec{\tau})$ will be a stable pair of transition probabilities $\vec{p}=(p_1,..,p_r)$ and transition times $\vec{\tau}=(\tau_1,...,\tau_r)$ as defined in the last chapter. To fix the scaling of the transition times once and for all, assume from now on that

$$(7.1)\qquad \sum_{i=1}^{r} m_i p_i t_i = \sum_{i=1}^{r} m_i p_i \int x d\tau_i(x)=1;$$

i.e., the average time it takes to get from one point in F to another is one (recall that m_i is the order of the equivalence class c_i defined at the beginning of Chapter V).

There must be many different ways of constructing Brownian motion on E; the method I shall use is based on nonstandard analysis, which has turned out to be a very convenient tool in limit constructions of this kind. The whole approach is closely related to Anderson's [2] nonstandard construction of ordinary Brownian motion (consult [1] for a closer description of this construction and also for an introduction to nonstandard probability theory). Toward the end of the chapter, I'll sketch a more traditional approach based on measure extension techniques.

As before

$$\hat{B}_n : [0,\infty)\times\Omega \to F^{(n)}$$

will be the random walks $(\vec{p},\vec{\tau})$ induce on $F^{(n)}$. Let $\bar{\Omega}={}^*\Omega$ be the nonstandard version of the sample space Ω, and for each probability measure P on Ω, let $\bar{P}$ denote the Loeb-measure of ${}^*P$ - hence $(\bar{\Omega},\bar{P})$ is an honest-to-goodness probability space. For each $n\in{}^*\mathbb{N}$, let $F^{(n)}={}^*\Psi^n(F)$ and note that since $F^{(n)}$ is finite for finite n, this notation agrees with earlier usage when $n\in\mathbb{N}$. Observe also that if n is infinite, then

$$E={}^\circ F^{(n)},$$

where $^\circ$ denotes the standard part operation. Abusing conventional notation slightly, I shall write $\{{}^*\hat{B}_n\}_{n\in{}^*\mathbb{N}}$ for the nonstandard version ${}^*(\{\hat{B}_n\}_{n\in\mathbb{N}})$ of the sequence $\{\hat{B}_n\}_{n\in\mathbb{N}}$. For each $n\in{}^*\mathbb{N}$, ${}^*\hat{B}_n$ is a map

$${}^*\hat{B}_n : {}^*[0,\infty)\times\bar{\Omega}\to F^{(n)}.$$

Observe that if n is finite and $x\in F^{(n)}$, then as a process on $(\bar{\Omega},\bar{P}^x)$, the standard part process $b_n:[0,\infty)\times\bar{\Omega}\to F^{(n)}$ defined by

$$b_n(t,\omega)={}^*\hat{B}_n(t,\omega)$$

is just a random walk started at x.

The idea is now to choose an infinitely large integer $N\in{}^*\mathbb{N}$ and then let

$$B:[0,\infty)\times\bar{\Omega}\to E$$

be the process defined by

$$B(t,\omega)={}^\circ({}^*\hat{B}_N(\lambda^N t,\omega))$$

(note that I'm rescaling time by the infinite factor λ^N). I shall prove that B is a continuous, strong Markov process with a law

that is independent of the integer N, and I shall refer to B as the Brownian motion on E induced by $(\vec{p},\vec{\tau})$. I would, of course, have liked to call it simply Brownian motion on E, but - at least as far as I know - there may be different stable pairs inducing different processes.

Before I prove that B is continuous, let me just remind you that in Proposition VI.7 I showed that there exist constants K and β such that

$$(7.2) \qquad P\{T_1<t\}<K^{-t^{-\beta}}$$

for all $t\in(0,1]$

VII.1 Proposition. For almost all $\omega\in\bar{\Omega}$, the path $B(\cdot,\omega)$ is continuous. In fact, if β satisfies (7.2), then there is a set $\Omega'\subset\bar{\Omega}$ of $\bar{P}$-measure one such that for each $\omega\in\Omega'$, there is a constant C_ω such that

$$|B(t)-B(s)| \leq C_\omega|t-s|^{\log\nu/\log\lambda}\left(\log\frac{1}{|t-s|}\right)^{\log\nu/\beta\log\lambda}$$

for all $s,t\in[0,1]$.

Proof: This argument is just an elaboration of the standard proof of Kolmogorov's Continuity Theorem. To simplify notation, I'll write

$$f(r)=|r|^{\log\nu/\log\lambda}\left(\log\frac{1}{|r|}\right)^{\log\nu/\beta\log\lambda}$$

and

$$X(t,\omega)=\hat{B}_N(\lambda^N t,\omega).$$

Let $a=\max\{|x-y|:x,y\in E\}$ be the diameter of E, and pick two arbitrary points $s,t,s<t$, in $[0,1]$. Note that if there are two points $u,v\in{}^*[s,t]$ such that $|X(u)-X(v)|>2a\nu^{-n}$ for some $n\in\mathbb{N}$, then X crosses an n-complex during the interval $[s,t]$.

According to (7.1) the probability for such a crossing is less than $K^{-(\lambda^n(t-s))^{-\beta}}$. Hence if C is a positive constant and $n\in\mathbb{N}$ is so large that $Cf(t-s)>2a\nu^{-n}$, then

$$^*P\{\exists u,v\in{}^*[s,t](|X(u)-X(v)|>Cf(t-s))\} < \exp[-(\lambda^n(t-s))^{-\beta}\log K]$$

Since

$$\lambda^{-n\beta}=(\nu^{-n})^{\beta\log\lambda/\log\nu} < \lfloor \frac{C}{2a} f(t-s)\rfloor^{\beta\log\lambda/\log\nu}$$

$$= C'\,(t-s)^{\beta}\,\log\frac{1}{(t-s)}\cdot(\log K)^{-1}$$

where $C' = (\frac{C}{2a})^{\beta\log\lambda/\log\nu}\log K$, it follows that

$$^*P\{\exists u,v\in{}^*[s,t](|X(u)-X(v)|>Cf(t-s))\}<|t-s|^{C'}.$$

As will become clear in a moment, it's essential to have $C'>1$, and I obtain this simply by choosing $C>2a(\log K)^{-\log\nu/\beta\log\lambda}$.

A spesial case of the inequality above is

$$^*P\{\exists u,v\in{}^*[\frac{k}{2^n},\frac{k+1}{2^n}](|X(u)-X(v)|>Cf(2^{-n}))\}<2^{-nC'}$$

for all pairs $(k,n)\in\mathbb{N}^2$ with $k<2^n$; by "overflow" the same formula holds for all pairs $(k,n)\in{}^*\mathbb{N}^2$ where n is less than some infinite $H\in{}^*\mathbb{N}$ and $k<2^n$. Thus for any given $m<H$

$$^*P\{\exists n(m<n<H)\exists k<2^n\exists u,v\in{}^*[\frac{k}{2^n},\frac{k+1}{2^n}](|X(u)-X(v)|>Cf(2^{-n}))\}$$

$$< \sum_{n=m}^{H}\sum_{k=0}^{2^n-1} 2^{-nC'} = \sum_{n=m}^{H} 2^{-n(C'-1)} < \frac{2^{-m(C'-1)}}{1-2^{-(C'-1)}}$$

which is infinitesimal when m is infinite. Consequently, there must be a set $\Omega'\subset\bar{\Omega}$ of $\bar{P}$-measure one such that for each $\omega\in\Omega'$, there is an integer $n(\omega)\in\mathbb{N}$ such that

$$\forall u,v\in{}^*[\frac{k}{2^n},\frac{k+1}{2^n}](|X(u)-X(v)|<Cf(2^{-n}))$$

for all $n, n(\omega)<n<H$, and all $k<2^n$.

It still remains to show that if $\omega\in\Omega'$, then there is a constant C_ω such that

$$|B(t)-B(s)| \leq C_\omega f(|t-s|)$$

for all $s,t\in[0,1]$. As a first preparatory step, I claim that there is a constant $D\in\mathbb{R}$ such that

$$\sum_{n=m}^{\infty} f(2^{-n}) \leq Df(2^{-m})$$

for all m. To prove this, first check that $\sum_{n=m}^{\infty} f(2^{-n})$ and $\int_m^{\infty} f(2^{-t})dt$ decrease at the same rate when $m\to\infty$, and then show that

$$\lim_{m\to\infty} \frac{\int_m^{\infty} f(2^{-t})dt}{f(2^{-m})}$$

is finite by using L'Hôpital's rule.

Here's the closing argument: Assume that $s,t\in[0,1]$ are given, and that $|s-t|<2^{-n(\omega)}$. Let $i\in\mathbb{N}$ be the integer such that $2^{-(i+1)}\leq|s-t|<2^{-i}$, and observe that there must be a $k\in\mathbb{N}$, $k<2^{-n}$, such that $\left|\frac{k}{2^i}-s\right|$ and $\left|\frac{k}{2^i}-t\right|$ are both less than 2^{-i}. But then s is of the form

$$s=\frac{k}{2^i}+\sum_{j=i+1}^{H}\frac{d_j}{2^j}$$

where d_j is 0,1 or -1 for $j<H$, and $|d_H|\leq 1$. Hence

$$|B(s)-B(\tfrac{k}{2^i})| \leq \sum_{p=i+1}^{H-1} |X(\tfrac{k}{2^i}+\sum_{j=i+1}^{p+1}\tfrac{d_j}{2^j}) - X(\tfrac{k}{2^i}+\sum_{j=i+1}^{p}\tfrac{d_j}{2^j})|$$

$$\leq \sum_{p=i+1}^{H-1} Cf(2^{-p}) \leq DCf(2^{-(i+1)}) \leq DCf(|t-s|),$$

and similarly $|B(t)-B(\frac{k}{2^i})| \leq DCf(|t-s|)$. This means that $|B(t)-B(s)| \leq 2DCf(|t-s|)$, provided that $|t-s|<2^{-n(\omega)}$, and the proposition is proved.

It follows from Proposition VI.7 that the conditions of the proposition above are satisfied when $\beta=\frac{\log q}{\log\lambda}$, and hence

$$|B(t)-B(s)| \leq C_\omega |t-s|^{\log\nu/\log\lambda}(\log\frac{1}{|t-s|})^{\log\nu/\log q}$$

for all $s,t\in[0,1]$. In this formula the exponent $\log\nu/\log\lambda$ of the $|t-s|$-term is clearly the best possible, but the exponent $\log\nu/\log q$ of the logarithmic term can almost certainly be improved; the estimates in the proof of Proposition VI.7 seem too crude to yield the best value for β. In any case, the argument above indicates that to get an exact expression for B's modulus of continuity, all one needs is a sharp estimate for the β in formula (7.2). I shall not pursue this question here, but only mention that using other methods, Barlow and Perkins [5] have shown that for the Sierpinski gasket the best value for β is $1-\log 2/\log 5$ (and in that case $\nu=2$ and $\lambda=5$!)

The next thing I would like to prove is that X is a strong Markov process. To do this I need an easy estimate and a few basic facts from potential theory. First the estimate:

VII.2 Lemma. For each $n\in\mathbb{N}$, let

$$\sigma_F^n(\omega)=\inf\{t:\hat{B}_n(t,\omega)\in F\}.$$

There is a constant C such that

$$E_n^x(\sigma_F^n) \leq C\lambda^n$$

for all $n\in N$ and all $x\in F^{(n)}$.

Proof: Assume first that $n=1$. By Proposition IV.11 any element $x\in F^{(1)}$ is connected to F, and hence $E_1^x(\sigma_F^1)$ is finite by elementary Markov chain theory. Let K be the maximum of the values $E_1^x(\sigma_F^1)$ for $x\in F^{(1)}$. I shall prove that the lemma holds with $C=\frac{K}{\lambda-1}$.

Assume now that $n\in\mathbb{N}$ and $x\in F^{(n)}$ are given. For each $k\leq n$, define a stopping time

$$\tau_k(\omega)=\inf\{t:\hat{B}_n(t,\omega) \text{ is a k-point}\}.$$

Clearly, $\tau_n=0$ and $\tau_0=\sigma_F^n$, and thus

$$\sigma_F^n=\sum_{k=1}^{n}(\tau_{k-1}-\tau_k)$$

By the definition of K and the stability of $(\vec{p},\vec{\tau})$

$$E_n^x(\tau_{k-1}-\tau_k)\leq K\lambda^{k-1},$$

and consequently

$$E_n^x(\sigma_F^n)\leq\sum_{k=1}^{n}K\lambda^{k-1}=K\frac{\lambda^n-1}{\lambda-1}\leq C\lambda^n,$$

which proves the lemma.

Note that together, Proposition VII.1 and Lemma VII.2 verify that λ^n is the right scaling factor for $\hat{B}_n$.

Let us turn to potential theory. Recall that if x and y are two elements in $F^{(n)}$, then $q_{x,y}^{(n)}$ is the probability that a particle which is in state x will next jump to state y; i.e.

$$q_{x,y}^{(n)}=P(B_n(i+1)=y \mid B_n(i)=x).$$

Note that $q_{x,y}^{(n)}=0$ unless x and y are n-neighbors. A function $u:F^{(n)}\to\mathbb{R}$ is n-harmonic if

$$u(x)=\sum_{y\in F^{(n)}} u(y)q_{x,y}^{(n)}$$

for all $x\in F^{(n)}-F$. If $\sigma_F^n(\omega)=\inf\{t:\hat{B}_n(t,\omega)\in F\}$, and f is any function $f:F\to\mathbb{R}$, then

$$u_f(x)=E_n^x\,(f(\hat{B}_n^x(\sigma_F^n))$$

is clearly an n-harmonic extension of f. It follows that an n-harmonic function always attains its maximum on F, and that two n-harmonic functions which agree on F agree everywhere. In fact, by Proposition IV.11 an n-harmonic function which attains its maximum at a point $x\in F^{(n)}-F$ is necessarily constant. All this is very easy and classical potential theory. The next lemma is just as easy, but more special to the situation we are dealing with here.

VII.3 Lemma. Let $n,m\in\mathbb{N}$, $n<m$, and assume that $u:F^{(n)}\to\mathbb{R}$ is n-harmonic and that $v:F^{(m)}\to\mathbb{R}$ is m-harmonic. If u and v agree on F, they also agree on $F^{(n)}$.

Proof: By Proposition V.6

$$u(x)=E_n^x(f(\hat{B}_n^x(\sigma_f^n)))=E_m^x(f(\hat{B}_m^x(\sigma_f^m)))=v(x)$$

for all $x\in F^{(n)}$, where $f=u\restriction F=v\restriction F$.

The next lemma shows that n-harmonic functions are very regular.

VII.4 Lemma. There exists a positive real number ε such that if u is n-harmonic and x,y belong to the same n-cell, then

$$|u(x)-u(y)|\le(1-\varepsilon)^n\,(\max_{r\in F}u(r)-\min_{r\in F}u(r)).$$

Proof: Let me first define ε. If

$$\sigma_F^1(\omega)=\inf\{t:\hat{B}_1(t,\omega)=F\},$$

choose $\varepsilon>0$ such that

$$\varepsilon<P_1^z\{\hat{B}_1(\sigma_F^1)=r\}<1-\varepsilon$$

for all $z\in F^{(1)}-F$, $r\in F$.

The proof goes by induction on n. Since the statement is trivially true for $n=0$, I need only consider the induction step from $n-1$ to n. Assume that x,y belong to the same n-cell $C=\psi_{i_1}\ldots\psi_{i_n}(F)$, and define

$$\hat{F}=\psi_{i_1}\circ\ldots\circ\psi_{i_{n-1}}(F)$$

$$\hat{F}^{(1)}=\psi_{i_1}\circ\ldots\circ\psi_{i_{n-1}}(F^{(1)})$$

$$\sigma_{\hat{F}}^n=\inf\{t:\hat{B}_n(t,\omega)\in\hat{F}\}.$$

Note that since $\hat{B}_n$ restricted to $\hat{F}^{(1)}$ is just a scaled copy of $\hat{B}_1$,

$$\varepsilon<P_n^z\{\hat{B}_n(\sigma_{\hat{F}}^n)=r\}<1-\varepsilon$$

for all $z\in\hat{F}^{(1)}-\hat{F}$, $r\in\hat{F}$. But then

$$\bar{a}\varepsilon+\underline{a}(1-\varepsilon)<u(z)<\bar{a}(1-\varepsilon)+\underline{a}\varepsilon$$

where $\bar{a}=\max_{r\in\hat{F}}u(r)$ and $\underline{a}=\min_{r\in\hat{F}}u(r)$. Since x and y belong to the same n-cell, at most one of them can be an element of $\hat{F}$ (recall Corollary IV.14), and hence it follows that

$$|u(x)-u(y)|<(1-\varepsilon)(\bar{a}-\underline{a}).$$

By Lemma VII.3, u is (n-1)-harmonic, and hence

$$\bar{a}-\underline{a}<(1-\varepsilon)^{n-1}(\max_{r\in F}u(r)-\min_{r\in F}u(r))$$

by the induction hypothesis. The lemma now follows from the last two inequalities.

As we shall soon see, the next proposition is just a nonstandard way of saying that B is a Feller process.

VII.5 Proposition. Assume that x and y are infinitely close elements in $F^{(N)}$, and that $f:E\to\mathbb{R}$ is continuous. If $X(t,\omega)={}^*\hat{B}_N(\lambda^N t,\omega)$, then for all finite $t\in{}^*[0,\infty)$

$$E_N^x({}^*f(X(t))) \approx E_N^y({}^*f(X(t))).$$

Proof: I can assume that x and y belong to the same H-complex $C=\phi_{i_1}\circ\ldots\circ\phi_{i_H}(E)$ for some infinite $H<N$, since if not there will be a third element $z\in F^{(N)}$ and complexes C', C'' of infinite order such that $x,z\in C'$ and $y,z\in C''$. Let H_0 be a smaller element in ${}^*\mathbb{N}-\mathbb{N}$ such that $H-H_0$ is infinite, and let $D=\phi_{i_1}\circ\ldots\circ\phi_{i_{H_0}}(E)$ be the H_0-complex x,y belong to. Define

$$\hat{F}=\phi_{i_1}\circ\ldots\circ\phi_{i_{H_0}}(F)$$

and let

$$\sigma_{\hat{F}} = \inf\{t:X(t,\omega)\in\hat{F}\}.$$

By Lemma VII.2, the expectations $E_N^x(\sigma_{\hat{F}})$ and $E_N^y(\sigma_{\hat{F}})$ are both of order of magnitude λ^{-H_0} and hence infinitesimal.

Let $r_1,\ldots,r_j$ be the elements in $\hat{F}$, and let

$$p_i^z={}^*P_N^z\{\sigma_{\hat{F}} = r_i\}.$$

The functions $z\to p_i^z$ are harmonic on $\phi_{i_1}\circ\ldots\circ\phi_{i_{H_0}}(\Psi^{H-H_0}(F))$, and thus

$$|p_i^x - p_i^y| < (1-\varepsilon)^{H-H_0} \tag{7.3}$$

by Lemma VII.4.

If $z \in F^{(N)}$, let m_z be the distribution on $F^{(N)}$ defined by

$$m_z(a) = \begin{cases} P_i^z & \text{if } a = r_i \text{ for some } i = 1, \ldots, j \\ 0 & \text{otherwise} \end{cases}$$

and, as usual, let ${}^*P^{m_z}$ be the probability measure governing ${}^*\hat{B}_N$ with initial distribution m_z. By (7.3)

$$E_N^{m_x}({}^*f(X(t))) \approx E_N^{m_y}({}^*f(X(t))),$$

and by definition of m_x, m_y

$$E_N^{m_x}({}^*f(X(t))) = E_N^x({}^*f(X(t+\sigma_{\hat{F}}))$$

$$E_N^{m_y}({}^*f(X(t))) = E_N^y({}^*f(X(t+\sigma_{\hat{F}}))).$$

The proposition now follows from the continuity of f and X, and the fact that $\sigma_{\hat{F}}$ is infinitesimal almost everywhere.

One of the small technical problems in proving that B is a Markov process comes from the fact that $\hat{B}_N$ is not Markov (since information about how long $\hat{B}_N$ has been waiting at a site tells us something about where it's likely to go next). The following lemma - which is just the law of large numbers in nonstandard disguise - will give us a simple way of avoiding this problem. I shall write ${}^*T_K$ for the K-th element in the sequence ${}^*(\{T_k\}_{k \in \mathbb{N}})$ (recall Conditions VI.1).

VII.6 Lemma. Let $K \in {}^*\mathbb{N}$ be such that $K \cdot \lambda^{-N}$ is finite. Then $\lambda^{-N}\,{}^*T_K(\omega) \approx \lambda^{-N}K$ for $\bar{P}$-a.a.ω.

Proof: Using that the random variables $({}^*\Delta T_i - 1)$, $i \in {}^*\mathbb{N}$, are

independent, have mean zero and finite variance σ^2, we get

$$E[(\lambda^{-N} {}^*T_K - \lambda^{-N}K)^2] = \lambda^{-2N}E[(\sum_{k=0}^{K-1} ({}^*\Delta T_k - 1))^2]$$

$$= \lambda^{-2N} \sum_{k=0}^{K-1} E[({}^*\Delta T_k - 1)^2] = \lambda^{-2N} \cdot K \cdot \sigma^2 \approx 0,$$

which proves the lemma.

Define a process $Y: {}^*[0,\infty) \times \bar{\Omega} \to F^{(N)}$ by

$$Y(t,\omega) = {}^*B_N([\lambda^N t], \omega),$$

where $[\ ]$ denotes integer part. By the lemma,

$$Y(t,\omega) \approx X(t,\omega)$$

for all finite t and almost all ω, and hence Proposition VII.5 holds with X replaced by Y.

Let $C(E)$ be the space of all continuous functions $f: E \to \mathbb{R}$ given the supremum norm. If $f \in C(E)$, $x \in E$ and $t \in [0,\infty)$, define

$$P^t f(x) = E_N^{\tilde{x}} ({}^*f(Y(t)))$$

where $\tilde{x}$ is any element in $F^{(N)}$ infinitely close to x. By Proposition VIII.5 (applied to Y instead of X) and the nonstandard characterization of continuity, $P^t f$ is continuous. I shall show that P^t is a Feller-Dynkin semigroup in the following sense (compare Williams [34, Definition 8.2]).

VII.7 Proposition. The P^t's are linear operators from $C(E)$ to $C(E)$ satisfying:

(i) If $f \in C(E)$ and $0 \le f(x) \le 1$ for all x, then $0 \le P^t f(x) \le 1$ for all x.

(ii) For all $s,t \in [0,\infty)$, $P^{t+s} = P^t \cdot P^s$, and $P^0 = 1$.

(iii) For all $f \in C(E)$, $\|P^t f - f\| \to 0$ as $t \downarrow 0$.

Proof: I have already checked that P^t maps $C(E)$ to $C(E)$, and part (i) is trivial from the definition of P^t. To check (ii), just observe that since Y is Markov

$$P^t \circ P^s f(x) \approx E_N^{\tilde{x}} (P^{s*} f(Y(t))) \approx E_N^{\tilde{x}} [E_N^{Y(t)} ({}^* f(Y(s)))]$$

$$= E_N^{\tilde{x}} ({}^* f(Y(t+s))) \approx P^{t+s} f(x).$$

Finally, to prove (iii) it suffices to show that for all $f \in C(E)$

$$P^t f(x) \to f(x) \quad \text{as} \quad t \downarrow 0$$

(see Williams [34, page 115]), and this follows immediately from the continuity of Y.

Any (right-) continuous process associated with a Feller-Dynkin semigroup is a strong Markov process, and hence it only remains to show that B is, indeed, associated with $\{P^t\}$, i.e.

$$\tilde{E}^{\tilde{x}} (f_1(B(t_1)) \cdot \ldots \cdot f_n(B(t_n)))$$

$$= (P^{t_1} f_1 P^{t_2 - t_1} f_2 \ldots . f_{n-1} P^{t_n - t_{n-1}} f_n)(x)$$

where $x = {}^\circ\tilde{x}$, for all increasing sequences $t_1 < \ldots < t_n$ and all function $f_1, .., f_n \in C(E)$.

This is quite easy. Write $Q^{t*} f$ for the function

$$Q^{t*} f(y) = E_N^y ({}^* f(Y(t))),$$

and note that by Proposition VII.5 and an easy induction argument,

$$(P^{t_1} f_1 P^{t_2 - t_1} f_2 \ldots . f_{n-1} P^{t_n - t_{n-1}} f_n)(x) \approx$$

$$\approx (Q^{t_1 *} f_1 Q^{t_2 - t_1 *} f_2 \ldots . {}^* f_{n-1} Q^{t_n - t_{n-1} *} f_n)(\tilde{x}).$$

Using the definition of B and the fact that Y is Markov, we get what we want:

$$\bar{E}^{\tilde{x}}\,(f_1(B(t_1))\cdot\ldots\cdot f_n(B(t_n)))$$

$$\approx E_N^{\tilde{x}}({}^*f_1(Y(t_1)).....{}^*f_n(Y(t_n)))$$

$$\approx (Q^{t_1}{}^*f_1Q^{t_2-t_1}{}^*f_2.....{}^*f_{n-1}Q^{t_n-t_{n-1}}{}^*f_n)(\tilde{x})$$

$$\approx (P^{t_1}f_1P^{t_2-t_1}f_2.....f_{n-1}P^{t_n-t_{n-1}}f_n)(x).$$

Let me collect the results above in the following theorem.

VII.8 Theorem. B is a strong Markov process with continuous paths.

It still remains to show that B (or rather B's law) is independent of the choice of the integer N. To this end, let H be another infinite integer, and define $B':[0,\infty)\times\bar{\Omega}\to E$ by

$$B'(t,\omega)= {}^\circ{}^*\hat{B}_H(\lambda^H t,\omega).$$

If $A_1, A_2,..,A_k$ are complexes of finite order, it is easy to see that for all $x\in F^{(\min(H,N))}$

$$\bar{P}_N^x\{B(t_1)\in A_1,....,B(t_k)\in A_k\}$$

$$\approx P_N^x\{{}^*\hat{B}_N(\lambda^N t_1)\in {}^*A_1,....,{}^*\hat{B}_N(\lambda^N t_k)\in {}^*A_k\}$$

$$\approx P_H^x\{{}^*\hat{B}_H(\lambda^H t_1)\in {}^*A_1,....,{}^*\hat{B}_H(\lambda^H t_k)\in {}^*A_k\}$$

$$\approx \bar{P}_H^x\{B'(t_1)\in A_1,....,B'(t_k)\in A_k\},$$

which shows that B and B' have the same finite dimensional distributions, and hence induce the same law on $C([0,\infty),E)$.

It is, of course, possible to turn the arguments above into an alternative construction of B which does not depend on non-

standard analysis; the idea is simply first to define the finite dimensional distributions

$$P^x_{t_1,..,t_k}(A_1\times\ldots\times A_k)=\lim_{n\to\infty} P^x_n\{\hat{B}_n(\lambda^n t_1)\in A_1,\ldots,\hat{B}_n(\lambda^n t_k)\in A_k\}$$

for each $x\in \bigcup_{n\in\mathbb{N}} F^{(n)}$, and then show that these finite dimensional distributions have a unique extension to a Borel measure. There are some not very difficult - but still rather annoying - difficulties to overcome, and I shall leave the details to the interested reader.

The following notation and terminology now seem natural. For each $x\in E$, let W_x be the Borel measure on $C([0,\infty),E)$ defined by

$$W_x(A)=\bar{P}^{\tilde{x}}_N\{\omega:B(\cdot,\omega)\in A\},$$

where $N\in {}^*\mathbb{N}-\mathbb{N}$ and where $\tilde{x}\in F^{(N)}$ is infinitely close to x. The measures $\{W_x\}_{x\in E}$ are called the <u>Wiener measures</u> induced by the stable point $(\vec{p},\vec{\tau})$. A process X with continuous paths is a <u>Brownian motion</u> induced by $(\vec{p},\vec{\tau})$ and starting at x if

$$W_x(A)=P\{\omega:X(\cdot,\omega)\in A\}$$

for all Borel sets A.

VIII. An invariance principle.

I argued in the Introduction that one of the main reasons for studying Brownian motion on fractals is to get a better understanding of the physics of fractal media. Having studied the construction above, and having, in particular, observed the extreme care with which I chose my transition probabilities and transition times, you have every reason to doubt that my processes have anything to do with physics - why should one expect Nature to choose her transition properties according to abstract fixed point theorems? The purpose of this section is to relieve your doubts by showing that the choice of basic transition properties is not as crucial as the previous sections may have led you to believe; in fact, a large variety of basic transition probabilities and transition times give rise to the same Brownian motion in the limit. More technically speaking, what I want to prove is a Donsker-type invariance principle for random walks and Brownian motions on fractals.

Let $(\vec{p},\vec{\tau})$ be a stable point, and let $\tilde{p}:\mathcal{P}\to\mathcal{P}$ be the map sending basic transition probabilities to composite transition probabilities as defined in Chapter V. A pair $(\vec{p}^0,\vec{\tau}^0)$ of transition probabilities and transition times belongs to $(\vec{p},\vec{\tau})$'s domain of attraction if

$$\sum_{i=1}^{r} m_i p_i^0 t_i^0 = \sum_{i=1}^{r} m_i p_i^0 \int x\, d\tau_i^0(x) = 1$$

and - more importantly - the iterates $\tilde{p}^n(\vec{p}^0)$ converge to $\vec{p}$ as n tends to infinity. As always, I'm assuming that $\vec{p}^0\in\mathcal{P}$ and that $\vec{\tau}^0$ has finite second moments.

Here's the nonstandard formulation of this chapter's main result.

VIII.1 Theorem. Assume that $(\vec{p}^0, \vec{\tau}^0)$ is in $(\vec{p}, \vec{\tau})$'s domain of attraction. Let $N \in {}^*\mathbb{N} - \mathbb{N}$ and assume that ${}^*\hat{b}_N$ is the random walk $(\vec{p}^0, \vec{\tau}^0)$ induces on $F^{(N)}$. Then there is a positive, nonstandard number D_N such that the standard process $b: [0,\infty) \times \bar{\Omega} \to E$ defined by

$$b(t,\omega) = {}^o({}^*\hat{b}_N(D_N t, \omega))$$

is a Brownian motion on E induced by $(\vec{p}, \vec{\tau})$.

From this nonstandard result, it's easy to deduce a standard theorem which just says that properly scaled versions of $\hat{b}_n$ (for finite n) converge weakly to Brownian motion induced by $(\vec{p}, \vec{\tau})$ (see Theorem VIII.4 below for the exact formulation).

Before I can prove the result above, I need to make my notation more explicit. Let $\vec{q} = (q_1, \ldots, q_r) \in \mathcal{P}$ be a set of basic transition probabilities. For any set $\vec{\sigma} = (\sigma_1, \ldots, \sigma_r)$ of basic transition times, let $\vec{\tilde{\sigma}} = (\tilde{\sigma}_1, \ldots, \tilde{\sigma}_r)$ be the composite transition times induced by $\vec{q}$ and $\vec{\sigma}$. I shall denote the map $\vec{\sigma} \to \vec{\tilde{\sigma}}$ by $\Theta_{\vec{q}}$, hence $\Theta_{\vec{q}}(\vec{\sigma}) = \vec{\tilde{\sigma}}$. If $\vec{s} = (s_1, .., s_r)$ and $\vec{\tilde{s}} = (\tilde{s}_1, .., \tilde{s}_r)$ are the expectations of $\vec{\sigma}$ and $\vec{\tilde{\sigma}}$, respectively, $T_{\vec{q}}$ will be the linear map sending $\vec{s}$ to $\vec{\tilde{s}}$.

Returning to the setting of the theorem, let $\vec{p}^n = \tilde{p}^{(n)}(\vec{p}^0)$ be the result of applying $\tilde{p}$ to $\vec{p}^0$ n times in succession. For simplicity, I shall write Θ_n and T_n for $\Theta_{\vec{p}^n}$ and $T_{\vec{p}^n}$, respectively, and I shall let Θ and T denote $\Theta_{\vec{p}}$ and $T_{\vec{p}}$.

For each n, let $\vec{\tau}^n$ be the transition times defined by

$$\vec{\tau}^n = \Theta_n \circ \Theta_{n-1} \circ \ldots \circ \Theta_1(\vec{\tau}^0).$$

To understand the probabilistic significance of $\vec{\tau}^n$, let σ_n be the stopping time

$$\sigma_n(\omega)=\inf\{t:\hat{b}_n(t,\omega)\in F-\{\hat{b}_n(0,\omega)\}\},$$

and choose $x,y\in F$ such that the pair (x,y) belongs to the equivalence class c_i. Then the i-th component τ_i^n of $\vec{\tau}^n$ is just the distribution of σ_n given that $\hat{b}_n(0)=x$ and $\hat{b}_n(\sigma_n)=y$.

Since $\vec{p}^n$ converges to $\vec{p}$, $\Theta_n(\vec{\sigma})$ converges to $\Theta(\vec{\sigma})$ for all $\vec{\sigma}$, as is easy to check. Moreover, the entries in T_n's matrix converge to the corresponding entries in T's matrix. Each T_n is a Perron-Frobenius map with a unique leading eigenvalue λ_n, and since the leading eigenvalue depends continuously on the entries of the matrix, the sequence $\{\lambda_n\}$ converges to the leading eigenvalue λ of T. Recall that λ is nothing but the time scaling factor associated with the stable point $(\vec{p},\vec{\tau})$.

It's often convenient to rescale $\vec{\tau}^n$ to get a distribution $\underline{\vec{\tau}}^n = (\underline{\tau}_1^n,..,\underline{\tau}_r^n)$ with total expectation

$$\Sigma m_i p_i^n \int x d\underline{\vec{\tau}}_i^n(x)=1.$$

I shall let $\hat{\Theta}_{n+1}$ be the map $\underline{\vec{\tau}}^n \to \underline{\vec{\tau}}^{n+1}$. If $\underline{\vec{t}}^n=(\underline{t}_1^n,..,\underline{t}_r^n)$ is the expectation of $\underline{\vec{\tau}}^n$, let $\hat{T}_{n+1}$ be the map sending $\underline{\vec{t}}^n$ to $\underline{\vec{t}}^{n+1}$. Assume that $\vec{t}$ is the expectation of the stable distribution $\vec{\tau}$, and let $\vec{s}$ be any r-dimensional vector with nonnegative components such that $\sum_{i=1}^{r} m_i p_i^0 s_i=1$. According to the Strong

Ergodicity Theorem (see Theorem 3.5 in Seneta [31], and, in particular, its Corollary), the scaled products

$$\hat{T}_n \circ \hat{T}_{n-1} \circ \ldots \circ \hat{T}_1(\vec{s}) \to \vec{t} \quad \text{as} \quad n \to \infty,$$

and the convergence is uniform in $\vec{s}$. Hence the expectations $\underline{\vec{t}}^n$ converge to $\vec{t}$ uniformly in the initial distribution $\vec{\tau}^0$.

The next step is to show that $\underline{\vec{\tau}}^n$ itself converges to $\vec{\tau}$. I shall make use of the following simple lemma:

VIII.2 Lemma. Let $\{x_n\}_{n\in\mathbb{N}}$ and $\{\varepsilon_n\}_{n\in\mathbb{N}}$ be two sequences of nonnegative real numbers, and assume that $\varepsilon_n \to 0$ as $n\to\infty$. If there is a real number $C<1$ such that

$$x_{n+1} \leq Cx_n + \varepsilon_n$$

for all $n\in\mathbb{N}$, then $x_n \to 0$ as $n\to\infty$.

Proof: By induction

$$x_{n+1} \leq C^n x_1 + \sum_{i=0}^{n-1} C^i \varepsilon_{n-i}.$$

Since the first term on the right hand side obviously tends to zero, it suffices to show that we can get $\sum_{i=0}^{n-1} C^i \varepsilon_{n-i}$ smaller than any given $\delta>0$ by choosing n sufficiently large.

First pick $n_0\in\mathbb{N}$ so large that

$$\sum_{i=n_0}^{\infty} C^i \max_{j\in\mathbb{N}} (\varepsilon_j) = \frac{C^{n_0}}{1-C} \max_{j\in\mathbb{N}} (\varepsilon_j) < \frac{\delta}{2},$$

and next choose $n>n_0$ such that

$$\max_{j>n-n_0} (\varepsilon_j) < \frac{\delta}{2}(1-C).$$

Then for any $m>n$

$$\sum_{i=0}^{m-1} C^i \varepsilon_{m-i} = \sum_{i=0}^{n_0-1} C^i \varepsilon_{m-i} + \sum_{i=n_0}^{m-1} C^i \varepsilon_{m-i}$$

$$< \max_{j>n-n_0} (\varepsilon_j) \sum_{i=0}^{n_0-1} C^i + \max_{j\in N} (\varepsilon_j) \sum_{i=n_0}^{m-1} C^i$$

$$< \max_{j>n-n_0} (\varepsilon_j) \frac{1}{1-C} + \max_{j\in N} (\varepsilon_j) \frac{C^{n_0}}{1-C}$$

$$< \frac{\delta}{2} + \frac{\delta}{2} = \delta,$$

and the lemma is proved.

Recall that in Chapter VI, I measured the distance between two distributions $\vec{\sigma}=(\sigma_1,..,\sigma_r)$ and $\vec{\sigma}'=(\sigma_1,...,\sigma_r')$ by the metric

$$d(\vec{\sigma},\vec{\sigma}') = (\sum_{i=1}^{r} d_i(\sigma_i,\sigma_i')^4)^{1/4},$$

where

$$d_i(\vec{\sigma}_i,\vec{\sigma}_i')= \inf\{ E[(S_i-S_i')^2]^{1/2}: S_i \text{ and } S_i' \text{ are random variables with distributions } \sigma_i \text{ and } \sigma_i', \text{ respectively}\}.$$

The idea is now to put $x_n=d(\vec{\tau},\underline{\vec{\tau}}^n)$ and show that the sequence $\{x_n\}$ satisfies an inequality $x_{n+1}<Cx_n+\varepsilon_n$ as in the lemma.

By the triangle inequality

$$x_{n+1}=d(\vec{\tau},\underline{\vec{\tau}}^{n+1})< d(\vec{\tau},\hat{\Theta}_{n+1}(\vec{\tau})) + d(\hat{\Theta}_{n+1}(\vec{\tau}), \hat{\Theta}_{n+1}(\underline{\vec{\tau}}^n))$$

Let me first check that the first term on the right hand side, $d(\vec{\tau},\hat{\Theta}_{n+1}(\vec{\tau}))$, tends to zero as n goes to infinity, and that it thus can be included in the ε_n-term in the lemma. To see this, observe that $\vec{\tau}$ is $\Theta(\vec{\tau})$ rescaled by λ and that $\hat{\Theta}_{n+1}(\vec{\tau})$ is $\Theta_{n+1}(\vec{\tau})$ rescaled by $\sum_{i=1}^{r} m_i p_i^{n+1} s_i^{n+1}$, where $\vec{s}^{n+1}=T_{n+1}(\vec{t}^n)$.

Since $\Theta_{n+1}(\vec{\tau}) \to \Theta(\vec{\tau})$, it thus suffices to show that

$\Sigma m_i p_i^{n+1} s_i^{n+1} \to \lambda$. But since $T_{n+1} \to T$ and $\vec{t}^n \to \vec{t}$,

$$\vec{s}^{n+1} = T_{n+1}(\vec{t}^n) \to T(\vec{t}) = \lambda \vec{t},$$

and consequently $\Sigma m_i p_i^{n+1} s_i^{n+1} \to \lambda \Sigma m_i p_i t_i = \lambda$.

So what about the second term in the inequality above - $d(\hat{\Theta}_{n+1}(\vec{\tau}), \hat{\Theta}_{n+1}(\vec{\underline{\tau}}^n))$? Let $\{S_k\}$ and $\{S_k'\}$ be two sequences of random variables satisfying Condition VI.1 with respect to $\vec{\tau}$ and $\vec{\underline{\tau}}^n$, respectively. If $q^n_{x,y}$ denotes the transition probabilities of the Markov chain $\vec{p}^n$ induces on $F^{(1)}$, then in the notation of Chapter VI

$$d(\hat{\Theta}_{n+1}(\vec{\tau}), \hat{\Theta}_{n+1}(\vec{\underline{\tau}}^n)) \leq$$

$$\leq \tilde{\lambda}_{n+1}^{-1} \Big(\sum_{i=1}^{r} (\tilde{Q}_i^{-1} \sum_{\vec{x} \in \Pi_i} E_{\vec{x}} \Big[\Big(\sum_{j=1}^{|\vec{x}|} (\Delta S_{j-1} - \Delta S'_{j-1}) \Big)^2 \Big] \prod_{j=1}^{|\vec{x}|} q^n_{x_{j-1} x_j} \Big)^2 \Big)^{1/4}$$

where $\tilde{\lambda}_{n+1} \equiv \Sigma m_i p_i^{n+1} s_i^{n+1}$. For each path $\vec{x}$

$$E_{\vec{x}} \Big[\Big(\sum_{j=1}^{|\vec{x}|} (\Delta S_{j-1} - \Delta S'_{j-1}) \Big)^2 \Big]$$

$$= \sum_{j=1}^{|\vec{x}|} E_{\vec{x}} \left[(\Delta S_{j-1} - \Delta S'_{j-1})^2 \right] + \sum_{j \neq k} E_{\vec{x}} (\Delta S_{j-1} - \Delta S'_{j-1}) E_{\vec{x}} (\Delta S_{k-1} - \Delta S'_{k-1})$$

$$= \sum_{j=1}^{|\vec{x}|} E_{\vec{x}} \left[(\Delta S_{j-1} - \Delta S'_{j-1})^2 \right] + \sum_{j \neq k} (t_{\ell_{j-1}} - \underline{t}^n_{\ell_{j-1}})(t_{\ell_{k-1}} - \underline{t}^n_{\ell_{k-1}})$$

where each ℓ_{j-1} is a number between 1 and r determined by the path $\vec{x}$. Since $t_{\ell_{j-1}} - \underline{t}^n_{\ell_{j-1}} \to 0$ as $n \to \infty$, it follows that

$$d(\hat{\theta}_{n+1}(\vec{\tau}),\ \hat{\theta}_{n+1}(\vec{\underline{\tau}}^n)) \leq \delta_n +$$

$$+ \tilde{\lambda}_{n+1}^{-1} \Big(\sum_{i=1}^{r} (\tilde{Q}_i^{-1} \sum_{\vec{x}\in\Pi_i} \sum_{j=1}^{|\vec{x}|} E_{\vec{x}} [(\Delta S_{j-1} - \Delta S'_{j-1})^2] \prod_{j=1}^{|\vec{x}|} q^n_{x_{j-1}x_j})^2 \Big)^{1/4}$$

where $\delta_n \to 0$ as $n\to\infty$ (Lemma V.3 guarantees that the infinite sum over all $\vec{x}\in\Pi_i$ doesn't cause any problems). Taking the infimum of the left hand side over all relevant sequences $\{S_k\}$ and $\{S'_k\}$, I get

$$d(\hat{\theta}_{n+1}(\vec{\tau}),\ \hat{\theta}_{n+1}(\vec{\underline{\tau}}^n)) \leq \delta_n +$$

$$+ \tilde{\lambda}_{n+1}^{-1} \Big(\sum_{i=1}^{r} (\tilde{Q}_i^{-1} \sum_{\vec{x}\in\Pi_i} \sum_{j=1}^{|\vec{x}|} d_{\ell_{j-1}}(\tau_{\ell_{j-1}}, \underline{\tau}^n_{\ell_{j-1}})^2 \prod_{j=1}^{|\vec{x}|} q^n_{x_{j-1}x_j})^2 \Big)^{1/4},$$

where ℓ_{j-1} is as above. Recall from Chapter VI that the i-th component of $T_{n+1}(\vec{y})$ for any vector $\vec{y}=(y_1,..,y_r)$ is

$$\tilde{Q}_i^{-1} \sum_{\vec{x}\in\Pi_i} \sum_{j=1}^{|\vec{x}|} y_{\ell_{j-1}} \prod_{j=1}^{|\vec{x}|} q^n_{x_{j-1}x_j}$$

(see (6.1)), and hence

$$d(\hat{\theta}_{n+1}(\vec{\tau}),\ \hat{\theta}_{n+1}(\vec{\underline{\tau}}^n)) \leq \delta_n + \tilde{\lambda}_{n+1}^{-1} \|T_{n+1}(d_1(\tau_1,\underline{\tau}^n_1)^2,..,d_r(\tau_r,\underline{\tau}^n_r)^2)\|^{\frac{1}{2}}$$

where $\|\cdot\|$ is the Euclidean norm in $\mathbb{R}^r$. From the proof of Theorem VI.5, I know that I can get $\|T^k\|/\lambda^k$ as close to one as I wish by choosing k large enough, and - since $\lambda>1$ - I can, in particular, get $\|T^k\|^{\frac{1}{2}}/\lambda^k<1$. In order not to make the notation unduly complicated, I shall assume that this inequality holds already for $k=1$; it's quite straightforward to extend the argument to cover the general case. But if now $\|T\|^{\frac{1}{2}}/\lambda<C<1$, then - since $\tilde{\lambda}_{n+1}\to\lambda$ and $T_{n+1}\to T$ - $\|T_{n+1}\|^{\frac{1}{2}}/\tilde{\lambda}_{n+1}<C$ for all sufficienty large n. For such n

$$d(\hat{\theta}_{n+1}(\vec{\tau}),\ \hat{\theta}_{n+1}(\vec{\underline{\tau}}^n)) \leq \delta_n + Cd(\vec{\tau},\vec{\underline{\tau}}^n).$$

Combining all the arguments above, we see that for sufficiently large n

$$d(\vec{\tau}, \underline{\vec{\tau}}^{n+1}) \leq Cd(\vec{\tau}, \underline{\vec{\tau}}^{n}) + \varepsilon_n,$$

where $\varepsilon_n = \delta_n + d(\vec{\tau}, \hat{\theta}_{n+1}(\vec{\tau}))$, and hence we can use Lemma VIII.2 to conclude:

VIII.3 Proposition. If $(\vec{p}^0, \vec{\tau}^0)$ is in $(\vec{p}, \vec{\tau})$'s domain of attraction, then $\underline{\vec{\tau}}^n \to \vec{\tau}$.

It is now quite easy to prove Theorem VIII.1. Let

$$D_n = \left(\sum_{i=1}^{r} m_i p_i^n t_i^n \right)$$

be the scaling factor which turns $\vec{\tau}^n$ into $\underline{\vec{\tau}}^n$, and note that for each $k \in \mathbb{N}$

$$\frac{D_n}{D_{n-k}} \to \lambda^k \quad \text{as} \quad n \to \infty.$$

Fix an infinite $N \in {}^*\mathbb{N}$, and let ${}^*\hat{b}_N$ be the random walk $(\vec{p}^0, \vec{\tau}^0)$ induces on $F^{(N)}$. If $C_N(t,\omega) = {}^*\hat{b}_N(D_N t, \omega)$ is a properly scaled version of ${}^*\hat{b}_N$, then for each $n \in \mathbb{N}$, the restriction of C_N to $F^{(n)}$ is a random walk with transition probabilities $\vec{p}^{N-n}$ and with transition times $\hat{\vec{\tau}}^{N-n}$, where $\hat{\vec{\tau}}^{N-n}$ is $\vec{\tau}^{N-n}$ rescaled by D_N. By assumption, $\vec{p}^{N-n}$ is infinitely close to $\vec{p}$, and by Proposition VIII.3 and the fact that $D_N/D_{N-n} \approx \lambda^n$, $\hat{\vec{\tau}}^{N-n}$ is infinitely close to $\vec{\tau}$ rescaled by λ^n. From this it follows easily that the standard process $b: [0,\infty) \times \bar{\Omega} \to E$ defined by $b(t,\omega) = {}^\circ C_N(t,\omega)$ has the same finite dimensional distributions as a Brownian motion induced by $(\vec{p}, \vec{\tau})$, and hence is a Brownian motion induced by $(\vec{p}, \vec{\tau})$. This proves

[. . . VIII.1]

Remark: It would be interesting to know if (or when) D_N is of order of magnitude λ^N. This seems to require information about how fast D_n/D_{n-1} converges to λ.

I have promised to give a standard translation of Theorem VIII.1 in terms of weak convergence. For a convenient formulation of the result, I'll first replace the discontinuous processes $\hat{b}_n$ by continuous processes β_n. Recall from Chapter VI that in order to define $\hat{b}_n$, I started with the Markov chain b_n induced by $\vec{p}^0$ and a sequence $\{T_k\}$ of random variables satisfying Condition VI.1 with respect to $\vec{\tau}^0$, and then let

$$\hat{b}_n(t,\omega)=b_n(k,\omega)$$

where k was the largest integer such that $T_k(\omega)<t$. The new process $\beta_n: [0,\infty)\times\Omega\to \mathbb{R}^k$ will be the continuous process obtained by interpolating linearly between consecutive values of $\hat{b}_n$, i.e.

$$\beta_n(t,\omega)=\hat{b}_n(T_k(\omega),\omega)+\frac{t-T_k(\omega)}{T_{k+1}(\omega)-T_k(\omega)}\cdot(\hat{b}_n(T_{k+1}(\omega),\omega)-\hat{b}_n(T_k(\omega),\omega))$$

where k is as above. It is now easy to show that β_n converges weakly to the Brownian motion induced by $(\vec{p},\vec{\tau})$ in the following sense.

VIII.4 Theorem. Let $\mathcal{C}$ be the space of all continuous functions $f:[0,\infty)\to\mathbb{R}$ with the topology of uniform convergence on compacts. Assume that $(\vec{p}^0,\vec{\tau}^0)$ is in the domain of attraction of the stable point $(\vec{p},\vec{\tau})$, and let B be the Brownian motion induced by $(\vec{p},\vec{\tau})$. Then there is a sequence $\{D_n\}$ of positive constants such

that

$$E_x(G(\beta_n(D_n\cdot,\omega))) \to E_x(G(B(\cdot,\omega))) \quad \text{as} \quad n\to\infty,$$

for all $x \in \bigcup_{n\in\mathbb{N}} F^{(n)}$ and all bounded, continuous functions $G: \zeta \to \mathbb{R}$.

Proof: Let $D_n = \left(\sum_{i=1}^{r} m_i p_i^n t_i^n\right)$ be as above. For each infinite $N \in \mathbb{N}$, the nonstandard process ${}^*\beta_N$ defines a standard process $y_N: [0,\infty)\times\bar{\Omega}\to E$ by

$$y_N(t,\omega) = {}^\circ({}^*\beta_N(D_N t,\omega)),$$

and according to Theorem VIII.1, y_N is a Brownian motion induced by $(\vec{p},\vec{\tau})$. For almost all ω, the path ${}^*\beta_N(D_N\cdot,\omega)$ is infinitely close (in ζ's topology) to the path $y_N(\cdot,\omega)$, and hence ${}^*G({}^*\beta_N(\cdot,\omega))$ is a lifting of $G(y_N(\cdot,\omega))$ in the sense of nonstandard measure theory. But then

$$E_x({}^*G({}^*\beta_N(D_N\cdot,\omega))) \approx \bar{E}_x(G(y_N(\cdot,\omega))) = E_x(G(B(\cdot,\omega)))$$

for all infinite N, and the theorem follows from the nonstandard characterization of convergence.

IX. The Laplacian and its eigenvalues.

Let E be a nested fractal constructed from a system $\Psi=\{\psi_1,\ldots,\psi_\mu\}$ of ν-similitudes, and let $(\vec{p},\vec{\tau})$ be a stable pair inducing a Brownian motion B on E. The Laplace operator associated with $(\vec{p},\vec{\tau})$ will simply be the infinitesimal generator of B. As I shall be primarily interested in the spectral properties of this Laplacian, I would like to regard it as a self-adjoint operator on a suitably chosen L^2-space. In probabilistic terms this is a question of constructing a measure m on E such that B is a symmetric Markov process with respect to m.

It's convenient first to take a look at what happens in the discrete setting. Fix an infinitely large, nonstandard integer N, and let ${}^*B_N$ be the Markov chain $\vec{p}$ induces on $F^{(N)}$. Let $\Delta t=\lambda^{-N}$, and introduce a hyperfinite timeline

$$T=\{k\Delta t: k\in {}^*\mathbb{N}_0\}.$$

The process $Y:T\times\Omega\to F^{(N)}$ is just the rescaled version of ${}^*B_N$ obtained by letting

$$Y(t,\omega)={}^*B_N\left(\frac{t}{\Delta t},\omega\right)$$

Recall that according to Lemma VII.6, Y is just the random walk ${}^*\hat{B}_N$ with an infinitesimal change of time. If $\{Q^t\}_{t\in T}$ denotes the semigroup generated by Y; i.e.

$$Q^t f(x)=E^x(f(Y(t)))$$

for all internal functions $f: F^{(N)}\to {}^*\mathbb{R}$, I shall now introduce an internal measure M on $F^{(N)}$ such that each Q^t becomes a symmetric operator with respect to the inner product

$$(f,g) = \sum_{x\in F^{(N)}} f(x)g(x)M(x).$$

Recall that for each $x\in F^{(N)}$, the multiplicity $\rho(x)$ is just the number of N-cells x belongs to. By Proposition IV.15, $\rho(x)$ is always finite. If A is an internal subset of $F^{(N)}$, let

$$M(A)= \frac{1}{\mu^N\cdot|F|} \sum_{x\in A} \rho(x),$$

where $|F|$ is the cardinality of F. Since $F^{(N)}$ consists of μ^N N-cells each with $|F|$ elements, it's clear that $M(F^{(N)})=1$.

Let $\{q_{x,y}\}_{x,y\in F^{(N)}}$ be the transition probabilities of Y (which are, of course, identical to the transition probabilities of ${}^*B_N$). By construction of Y and M, we get immediately that

$$q_{x,y}\, M(x) = q_{y,x}\, M(y)$$

for all $x,y\in F^{(N)}$. This simple observation is crucial; not only does it show that M is an invariant measure for Y, but it also implies that $Q^{\Delta t}$ is symmetric:

$$(Q^{\Delta t}f,g) = \sum_{x\in F^{(N)}} \Big(\sum_{y\in F^{(N)}} f(y)q_{x,y}\Big)g(x)M(x)=$$

$$\sum_{x,y\in F^{(N)}} f(y)g(x)q_{x,y}M(x) = \sum_{x,y\in F^{(N)}} f(y)g(x)q_{y,x}M(y) =$$

$$\sum_{y\in F^{(N)}} f(y)\Big(\sum_{x\in F^{(N)}} g(x)q_{y,x}\Big)M(y) = (f,Q^{\Delta t}g).$$

Since $\{Q^t\}$ is a semigroup, this means that Q^t is symmetric for all $t\in T$.

It's now quite easy to translate these results into standard terms. Let $\bar{M}$ be the Loeb-measure of M, and define a measure m on E by

$$m(A)=\bar{M}(st^{-1}(A)),$$

where st is the standard part map in $\mathbb{R}^k$. By elementary

Loeb-measure theory (see, e.g., Section 3.4 in [1]), m is a Radon probability measure on E.

Assume now that Y has initial distribution M, and let $B:[0,\infty)\times\bar{\Omega}\to E$ be defined by

$$B(t,\omega)= {}^{\circ}Y(\bar{t},\omega)$$

where $\bar{t}$ is the largest element in T smaller than or equal to t. Clearly, B is a Brownian motion with initial distribution m. Note that since M is an invariant distribution for Y, m must be an invariant distribution for B.

Given a $t\in[0,\infty)$ and a Borel function $f:E\to\mathbb{R}$, define

$$T^t f(x)=E^x(f(B(t))).$$

IX.1 Lemma. Each T^t is a bounded operator from $L^2(E,m)$ to $L^2(E,m)$ with norm 1.

Proof: Assume first that f is a nonnegative, bounded function. Then $T^t f$ is also bounded, and since m is an invariant distribution for B:

$$(9.1)\quad \int f(x)^2 dm(x) = E(f(B(t))^2)= \int E^x(f(B(t))^2)dm(x) \geq \int E^x(f(B(t)))^2 dm(x)= \int T^t f(x)^2 dm(x).$$

Let now f be an arbitrary, nonnegative function in $L^2(E,m)$, and let $f_n=f\wedge n$ for each $n\in\mathbb{N}$. Since

$$\int (T^t f(x)-T^t f_n(x))dm(x) = \int E^x(f(B(t))-f_n(B(t)))dm(x)= E(f(B(t))-f_n(B(t))) = \int (f(x)-f_n(x))dm(x)\to 0,$$

the sequence $\{T^t f_n(x)\}$ increases to $T^t f(x)$ for almost all x, and hence

$$\int f(x)^2 dm(x) = \lim_{n\to\infty} \int f_n(x)^2 dm(x) \geq \lim_{n\to\infty} \int Tf_n(x)^2 dm(x) = \int T^t f(x)^2 dm(x)$$

by (9.1) and the Monotone Convergence Theorem. To extend this result to arbitrary elements of $L^2(E,m)$, just note that

$$\|f\| = \||f|\| \geq \|T^t|f|\| \geq \|T^t f\|$$

Hence T^t is a bounded operator with $\|T^t\| \leq 1$. But since $T^t 1 = 1$, the norm must be exactly one.

IX.2 Proposition: $\{T^t\}_{t\in[0,\infty)}$ is a strongly continuous, contraction semigroup of symmetric operators on $L^2(E,m)$.

Proof: Let me first show that T^t is symmetric. According to Proposition VII.7, T^t maps continuous functions to continuous functions, and hence $Q^{\bar{t}\star} f(x) \approx T^t f(x)$ for all continuous f and all $x \in F^{(N)}$. But then

$$\int T^t f(x) g(x) dm(x) = \circ \sum_{y\in F^{(N)}} Q^{\bar{t}\star} f(y)^* g(y) M(y)$$

$$= \circ \sum_{y\in F^{(N)}} {}^* f(y) Q^{\bar{t}\star} g(y) M(y) = \int f(x) T^t g(x) dm(x)$$

for all continuous f and g, and hence T^t is symmetric.

Recall that $\{T^t\}$ is strongly continuous if for all $f \in L^2(E,m)$,

$$\int (f - T^t f)^2 dm \to 0$$

as $t \downarrow 0$. By Proposition VII.7, $T^t f$ converges uniformly to f for all continuous f, and hence strong continuity follows by a simple density argument. Since the lemma shows that $\{T^t\}$ is a contraction semigroup, the proof is complete.

I can now define the Laplace operator Δ associated with the stable point $(\vec{p},\vec{\tau})$ to be the infinitesimal generator of the semigroup $\{T^t\}$; i.e., Δ is the operator defined on the domain

$$\mathfrak{D}[\Delta]=\{f\in L^2(E,m):\lim_{t\downarrow 0}\frac{T^tf-f}{t} \text{ exists in the strong sense}\}$$ by

$$\Delta f=\lim_{t\downarrow 0}\frac{T^tf-f}{t}\ .$$

IX.3 Theorem. The Laplacian Δ is a self-adjoint nonpositively definite operator on $L^2(E,m)$

Proof: By general operator theory (see e.g. Lemma 1.3.1 in Fukushima [14]) any strongly continuous contraction semigroup of symmetric operators has a nonpositively definite, self-adjoint infinitesimal generator.

In the rest of this chapter, I shall study the asymptotic distribution of the eigenvalues of Δ. If

$n(\alpha)$= the number of eigenvalues of $-\Delta$ less than or equal to α (counted with multiplicity), I want to show that there are positive constants $k,c\in\mathbb{R}$ such that

$$k\alpha^{\log\mu/\log\lambda}\leq n(\alpha)\leq c\alpha^{\log\mu/\log\lambda}.$$

When I presented the basic ideas of the proof at the end of Chapter III, the trace-functions $Tr_N(t)$ played an essential part, and when I now want to make the argument rigorous, my first task is to control the behavior of these functions. The estimates I shall use are based on very similar estimates in Barlow and Perkins [5].

Let us begin with a few definitions. Choose $z\in F$ and define a stopping time τ by

$$\tau(\omega)=\min\{n:B_1^z(n,\omega)\in F-\{z\}\}$$

(where, as usual, B_1 is the Markov chain on $F^{(1)}$ induced by the basic transition probabilities $\vec{p}=(p_1,..,p_r)$). Let

$$N(\omega)=\#\{k<\tau(\omega):B_1^z(k,\omega)=z\},$$

and define

$$\eta=E^z(N(\omega)).$$

By symmetry considerations, η is independent of z, and - as we shall soon seen - it is actually equal to λ/μ.

For each $n\in\mathbb{N}$, let

$$\sigma_n(\omega)=\min\{k:B_n(k,\omega)\in F\},$$

and define the "Green function" $g_n:F^{(n)}\times F^{(n)}\to\mathbb{R}$ by

$$g_n(x,y)=E^x\left(\sum_{k=0}^{\sigma_n-1} 1_{\{y\}}(B_n(k,\omega))\right).$$

If $\rho(x)$ is the multiplicity of the state x (see the definition at the end of Chapter IV), then clearly

$$g_n(x,y)\rho(x)=g_n(y,x)\rho(y),$$

and hence the "normalized" function

$$\tilde{g}_n(x,y) = \frac{g_n(x,y)}{\rho(y)}$$

is symmetric in x and y. I want to find upper and lower bounds for $\tilde{g}_n$.

Notice first that by definition of η

(9.2) $\tilde{g}_{n+1}(x,y)=\eta\tilde{g}_n(x,y)$ whenever $x,y\in F^{(n)}$.

Assume next that $x,y\in F^{(n+1)}$ belong to two different n-complexes, and let $x_1,x_2,..,x_\xi$ and $y_1,y_2,..,y_\xi$ be the elements in the corresponding n-cells (see Figure 14). Let

$$\tau_x(\omega)=\min\{k:B^x_{n+1}(k,\omega)=x_1,x_2,..,\text{ or } x_\xi\}$$

$$\tau_y(\omega)=\min\{k:B^y_{n+1}(k,\omega)=y_1,y_2,..,\text{ or } y_\xi\}$$

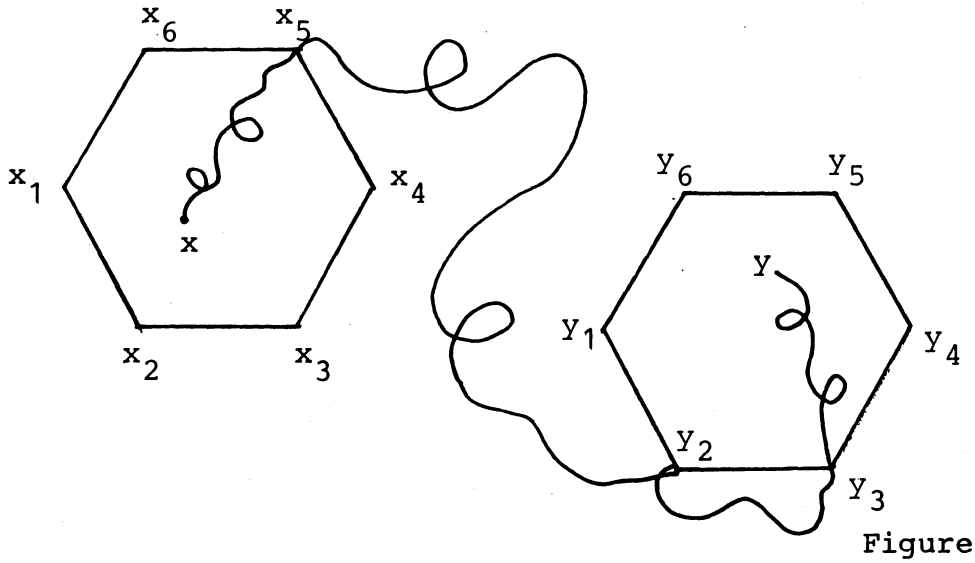


Figure 14

and define

$$p_i=P^x(B^x_{n+1}(\tau_x)=x_i)$$

$$q_j=P^y(B^y_{n+1}(\tau_y)=y_j)$$

Since a path starting at x has to pass through either $x_1,x_2,..,$ or x_ξ before it can reach y, we get

$$\tilde{g}_{n+1}(x,y)=\sum_{i=1}^{\xi}\tilde{g}_{n+1}(x_i,y)p_i$$

By symmetry, $\tilde{g}_{n+1}(x_i,y)=\tilde{g}_{n+1}(y,x_i)$, and since a path starting at y has to hit $y_1,..,y_\xi$ before it hits x_i, we see that

$$\tilde{g}_{n+1}(y,x_i)=\sum_{j=1}^{\xi}\tilde{g}_{n+1}(y_j,x_i)q_j$$

Hence

$$\tilde{g}_{n+1}(x,y)=\sum_{i,j=1}^{\xi}\tilde{g}_{n+1}(x_i,y_j)p_iq_j$$

or - using (9.2) -

$$(9.3)\qquad \tilde{g}_{n+1}(x,y)=\eta \sum_{i,j=1}^{\xi} \tilde{g}_n(x_i,y_j)p_iq_j$$

It is easy to see that there must be two different 2-complexes C_1 and C_2 which do not contain any element from F. For each $n\in^* \mathbb{N}$, let C_1^n and C_2^n be the set of all n-points belonging to C_1 and C_2, respectively. Let

$$k=\min\{\tilde{g}_2(x,y): x\in C_1^2,\ y\in C_2^2\}$$

$$K=\max\{\tilde{g}_2(x,y): x\in C_1^2,\ y\in C_2^2\}.$$

Since C_1 and C_2 do not intersect F, it follows from Proposition IV.11 that k is strictly positive. By (9.3) and induction,

$$k\eta^{(n-2)} \leq \tilde{g}_n(x,y) \leq K\eta^{(n-2)}$$

for all $x\in C_1^n$, $y\in C_2^n$. On the other hand, for each fixed $x\in C_1^n$, the sum

$$\sum_{y\in C_2^n} \tilde{g}_n(x,y)\lambda^{-n}$$

must be noninfinitesimal and finite even for infinite n, since it just measures the average time a particle starting in x spends in C_2 before it hits F. As the number of elements in C_2^n is of the order of magnitude μ^n, this means that $\eta^{(n-2)}\lambda^{-n}\mu^n$ must be finite and noninfinitesimal for all $n\in^* \mathbb{N}$. But this is only possible if $\eta=\lambda/\mu$. Hence we have proved:

IX.4 Proposition: $\eta=\frac{\lambda}{\mu}$.

By definition, η is larger than one, and thus we get

IX.5 Corollary. $\lambda>\mu$.

Finally, combining Proposition IX.4 and formula (9.3), we see that

IX.6 Proposition. If $x,y\in F^{(n+1)}$ belong to different n-complexes, then

$$\tilde{g}_{n+1}(x,y)=\frac{\lambda}{\mu}\sum_{i,j=1}^{\xi}\tilde{g}_n(x_i,y_j)p_iq_j,$$

where x_i,y_j,p_i,q_j are as above.

So far we have only estimated $\tilde{g}_n(x,y)$ when x and y are apart, but to deal with the trace-functions we need to know what happens when $x=y$. Here's the first result in that direction:

IX.7 Lemma. There is a constant $C\in\mathbb{R}_+$ such that if $x,y\in F^{(n+1)}$ belong to the same n-complex, then

$$\tilde{g}_{n+1}(x,y)\leq C+\frac{\lambda}{\mu}\sum_{i,j=1}^{\xi}\tilde{g}_n(x_i,x_j)p_iq_j.$$

(The notation is as before, but since x and y now belong to the same n-complex, $x_1=y_1$, $x_2=y_2,\ldots,x_\xi=y_\xi$).

Proof: Let

$$C=\max\{g_1(x,y):x\in F^{(1)},\ y\in F^{(1)}\},$$

and define

$$\tau_{n+1}(\omega)=\min\{k:B_{n+1}(k,\omega)=x_1,x_2,\ldots,\text{ or } x_\xi\}$$

Clearly,

$$\tilde{g}_{n+1}(x,y)=\frac{1}{\rho(y)}\Big[\sum_{k=0}^{\tau_{n+1}-1}1_{\{y\}}(B_{n+1}(k,\omega))+\sum_{k=\tau_{n+1}}^{\sigma_{n+1}-1}1_{\{y\}}(B_{n+1}(k,\omega))\Big]$$

(recall that $\sigma_{n+1}(\omega)=\min\{k:B_{n+1}(k,\omega)\in F\}$). The first term in this sum cannot be larger than C, while the second term equals

$$\sum_{i=1}^{\xi} \tilde{g}_{n+1}(x_i,y)p_i,$$

and hence

$$\tilde{g}_{n+1}(x,y) \leq C + \sum_{i=1}^{\xi} \tilde{g}_{n+1}(x_i,y)p_i.$$

By essentially the same argument,

$$\tilde{g}_{n+1}(x_i,y) = \tilde{g}_{n+1}(y,x_i) = \sum_{j=1}^{\xi} \tilde{g}_{n+1}(x_j,x_i)q_j$$

and hence

$$\tilde{g}_{n+1}(x,y) \leq C + \sum_{i,j=1}^{\xi} \tilde{g}_{n+1}(x_i,x_j)p_i q_j$$

$$= C + \frac{\lambda}{\mu} \sum_{i,j=1}^{\xi} \tilde{g}_n(x_i,x_j)p_i q_j,$$

which proves the lemma.

Combining the last two result, we get:

IX.8 Proposition. There is a constant $K \in \mathbb{R}_+$ such that

$$\tilde{g}_n(x,y) \leq K\left(\frac{\lambda}{\mu}\right)^n$$

for all $n \in {}^* \mathbb{N}$ and all $x,y \in F^{(n)}$.

Proof: Let

$$v_k = \max\{\tilde{g}_k(x,y) : x,y \in F^{(k)}\}.$$

By the last two results there is a constant C such that

$$v_{k+1} \leq C + \frac{\lambda}{\mu} v_k,$$

and hence by induction

$$v_n \leq C \sum_{i=0}^{n-1} \left(\frac{\lambda}{\mu}\right)^i + \left(\frac{\lambda}{\mu}\right)^n v_0.$$

Summing the geometric series, we get

$$v_n \leq \frac{C}{\frac{\lambda}{\mu} - 1}\left[\left(\frac{\lambda}{\mu}\right)^n - 1\right] + \left(\frac{\lambda}{\mu}\right)^n v_0 \leq K\left(\frac{\lambda}{\mu}\right)^n,$$

where $K=\frac{C}{\frac{\lambda}{\mu}-1}+v_0$.

The results so far give absolutely no information about what happens if the process starts at a point in F, but this is easily remedied:

IX.9 Lemma. Assume that $x\in F$ and let

$$\sigma_n(\omega)=\min\{k:B_n(k,\omega)\in F-\{x\}\}.$$

Then for all $y\in F^{(n)}$,

$$E^x\Big(\sum_{k=0}^{\sigma_n-1} 1_{\{y\}}(B_n(k,\omega))\Big)\leq\rho(y)\Big(\frac{\lambda}{\mu}\Big)^n,$$

where $\rho(y)$ is the multiplicity of y.

Proof: If $y=x$, then

$$E^x\Big(\sum_{k=0}^{\sigma_n-1} 1_{\{x\}}(B_n(k,\omega))\Big)=\eta^n=\Big(\frac{\lambda}{\mu}\Big)^n$$

by the definition of η and Proposition IX.4. If $y\neq x$, then

$$E^x\Big(\sum_{k=0}^{\sigma_n-1} 1_{\{y\}}(B_n(k,\omega))\Big)=\rho(y)E^y\Big(\sum_{k=0}^{\sigma_n-1} 1_{\{x\}}(B_n(k,\omega))\Big)$$

$$\leq\rho(y)E^x\Big(\sum_{k=0}^{\sigma_n-1} 1_{\{x\}}(B_n(k,\omega))\Big)\leq\rho(y)\Big(\frac{\lambda}{\mu}\Big)^n,$$

which proves the lemma.

We are now ready to begin our study of trace-functions. Let

$$Y_N(t,\omega)=B_N([\lambda^N t],\omega)$$

be the rescaled version of B_N, and let

$$\sigma(\omega)=\min\{t:Y_N(t,\omega)\in F\}.$$

Add a cemetery state ∞ to the state space $F^{(N)}$, and let $\tilde{Y}_N$ be the stopped process

$$\tilde{Y}_N(t,\omega)=\begin{cases} Y_N(t,\omega) & \text{if } t<\sigma_1(\omega) \\ \infty & \text{otherwise.}\end{cases}$$

Note that $\tilde{Y}_N$ is a symmetric Markov process (with respect to the measure M introduced at the beginning of the chapter) taking values in $(F^N-F)\cup\{\infty\}$. We shall study the semigroups

$$Q_N^t f(x)=E^x[f(Y_N(t))]$$

$$\tilde{Q}_N^t f(x)=E^x[f(\tilde{Y}_N(t))]$$

defined for all $t\in\{k\cdot\lambda^{-N}:k\in^* \mathbb{N}_0\}$. As usual the <u>trace of</u> Q_N^t is defined by

$$Tr_N(t)=\sum_{i=1}^{H}(Q_N^t f_i,f_i),$$

where $\{f_i\}$ is any orthonormal basis for $L^2(F^{(N)},M)$, and similarly the <u>trace of</u> $\tilde{Q}_N^t$ is given by

$$\tilde{Tr}_N(t)=\sum_{i=1}^{H-\xi}(\tilde{Q}_N^t \tilde{f}_i,\tilde{f}_i),$$

where $\{\tilde{f}_i\}$ is any orthonormal basis for $L^2(F^{(N)}-F,M)$.

Let $x_1,x_2,..,x_H$ be an enumeration of the points in $F^{(N)}$ which ends with the ξ elements in F, and define f_i by

$$f_i(x_j)=\begin{cases} M(x_i)^{-\frac{1}{2}} & \text{if } i=j \\ 0 & \text{otherwise.}\end{cases}$$

Then $\{f_i\}$ is an orthonormal basis, and

$$Tr_N(t)=\sum_{i=1}^{H}(Q_N^t f_i,f_i)=\sum_{i=1}^{H}q_N^t(x_i,x_i)$$

$$\tilde{T}r_N(t)=\sum_{i=1}^{H-\xi}\tilde{q}_N^t(x_i,x_i),$$

where

$$q_N^t(x,x)=E^x(Y_N(t)=x)$$

$$\tilde{q}_N^t(x,x)=E^x(\tilde{Y}_N(t)=x).$$

Another way of looking at $Tr_N(t)$ is to start with the eigenvalues $\{\alpha_i\}_{i=1}^H$ of the discrete infinitesimal generator

$$A_N=\frac{Q_N^{\Delta t}-I}{\Delta t}$$

(recall that $\Delta t=\lambda^{-N}$), and observe that

$$Tr_N(t)=\sum_{i=1}^{H}(1+\alpha_i\Delta t)^{t/\Delta t}.$$

All the eigenvalues α_i are negative, and it is easy to check that $|1+\alpha_i\Delta t|\leq 1$. This means that as long as we restrict ourselves to even values $0,2\Delta t,4\Delta t,\ldots$ of t, the function $t\to Tr_N(t)$ will be decreasing. The same observation holds, of course, for $\tilde{T}r_N(\cdot)$.

IX.10 Proposition. $Tr_N(t)$ and $\tilde{T}r_N(t)$ are finite for all noninfinitesimal $t\in{}^*\mathbb{R}$ and all $N\in{}^*\mathbb{N}$.

Proof: Since $\tilde{T}r_N(t)\leq Tr_N(t)$, it suffices to show that $Tr_N(t)$ is finite. Define a sequence $\{\sigma_k\}$ of stopping times by

$$\sigma_1(\omega)=\min\{t:Y_N(\omega,t)\in F\}$$

and - assuming that σ_{k-1} is already defined -

$$\sigma_k(\omega)=\min\{t>\sigma_{k-1}(\omega):Y_N(t,\omega)\in F-\{Y(\sigma_{k-1})\}\}.$$

Clearly,

$$\sum_{s=0}^{t} q_N^t(x,x)\lambda^{-N} \leq g_n(x,x)\lambda^{-N} +$$

$$\sum_{k=1}^{\infty} E^{\sigma_k}\Big[\sum_{s=\sigma_k}^{\sigma_{k+1}-1} 1_{\{x\}}(Y_N(t))\cdot 1_{\{\sigma_k<t\}}\Big]\lambda^{-N}$$

By elementary Markov chain theory, there are real constants D and ζ, $0<\zeta<1$, such that

$$P\{\sigma_k<t\} \leq D\zeta^k$$

for all $k\in^* \mathbb{N}$. By Lemma IX.9 and Proposition IX.8, we thus get

$$\sum_{s=0}^{t} q_N^t(x,x)\lambda^{-N} \leq \rho(x)\cdot K\cdot\mu^{-N} + \sum_{k=1}^{\infty} \rho(x)\mu^{-N}D\zeta^k \leq \hat{C}\,\mu^{-N}$$

for some real constant $\hat{C}$. Summing over all $x\in F^{(N)}$, we see that

$$\sum_{s=0}^{t} Tr_N(t)\lambda^{-N} \leq \hat{C}\,.$$

Since $Tr_N(t)$ is (essentially) decreasing, the proposition follows.

I now have the estimates I shall need. My next task is to show that if N is infinite, then trace(T^t) equals the standard part of $Tr_N(t)$; this is just a nonstandard way of saying that $Tr_n(t)$ converges to trace(T^t) when $n\in\mathbb{N}$ goes to infinity. The next three results deal with this question, but before I turn to them, I need a few definitions. Just as A_N denotes the discrete infinitesimal generator of Y_N, $\tilde{A}_N$ will denote the discrete infinitesimal generator of the stopped process $\tilde{Y}_N$, i.e.

$$\tilde{A}_N = \frac{\tilde{Q}_N^{\Delta t} - I}{\Delta t}\,.$$

It's easy to check that the standard part of $\tilde{Y}_N$ is just B stopped the first time it hits F; the semigroup associated with this process will be called $\{\tilde{T}^t\}$.

IX.11 Lemma. Assume that N is infinite and t noninfinitesimal. Then

$$^\circ Tr_N(t)=\Sigma\{\,^\circ(1+\alpha\Delta t)^{t/\Delta t}: \alpha \text{ is a finite eigenvalue of } A_N\}.$$

The corresponding result holds for $\tilde{Tr}_N$.

Proof: If not, there would be an infinite H such that

$$\Sigma\{(1+\alpha\Delta t)^{t/\Delta t}: \alpha>H \text{ is an eigenvalue of } A_N\}$$

was not infinitesimal. But then

$$^\circ Tr_N(t/2)>\,^\circ\Sigma\{(1+\alpha\Delta t)^{t/2\Delta t}: \alpha>H \text{ is an eigenvalue}\}=\infty$$

contradicting the proposition above. The same argument works for $\tilde{Tr}_N$.

IX.12 Proposition. Assume that v is an eigenfunction of A_N corresponding to a finite eigenvalue α. Then v is S-bounded and S-continuous. The corresponding result holds for $\tilde{A}_N$.

Proof: That v is S-bounded means that it has a finite maximum. Assume not, and let x be a point where it attains its infinite maximum. It is easy to see that there must be an infinite $K\in{}^*\mathbb{N}$ and a point $y\in F^{(N)}$ such that:

(i) x and y belong to the same K-complex $\hat{C}$

(ii) $|x-y|$ is infinitesimal compared to ν^{-K},

(iii) $v(y)\leq\frac{1}{2}v(x)$.

Let σ be the first time Y_N hits one of the vertices of $\hat{C}$ (a vertex is just an element in the corresponding K-cell C). By Lemma VII.2, σ is infinitesimal $\bar{P}^x$- and $\bar{P}^y$- a.s., and by (ii)

and Lemma VII.4, $Y_N(\sigma)$ has almost the same distribution with respect to P^x and P^y. Since v is an eigenfunction with eigenvalue α, then for $z=x$ or $z=y$,

$$v(z)=-\alpha \sum_{k=0}^{\infty} Q^{k\cdot\Delta t}v(z)\Delta t = -\alpha E^z\Big(\sum_{k=0}^{\infty} v(Y(k\Delta t))\Delta t\Big)$$

$$= -\alpha E^z\Big(\sum_{k=0}^{\sigma/\Delta t-1} v(Y(k\Delta t))\Delta t\Big) - \alpha E^z\Big(\sum_{\sigma/\Delta t}^{\infty} v(Y(k\Delta t))\Delta t\Big).$$

Note that since σ is infinitesimal, the first term on the right hand side is infinitesimal compared to $v(x)$. Note also that since $Y(\sigma)$ has almost the same distribution with respect to P^x and P^y, we must have

$$(9.4)\qquad \frac{E^y\Big(\sum_{k=\sigma/\Delta t}^{\infty} v(Y_N(k\Delta t))\Delta t\Big)}{E^x\Big(\sum_{k=\sigma/\Delta t}^{\infty} v(Y_N(k\Delta t))\Delta t\Big)} \approx 1$$

Combining these two observations, we get $\frac{v(y)}{v(x)} \approx 1$, which contradicts (iii)

To show that v is S-continuous, we must start with two infinitely close elements $x,y\in F^{(N)}$ and prove that $v(x)\approx v(y)$. Choose an infinite $K\in^* \mathbb{N}$ such that $|x-y|$ is infinitesimal compared to ν^{-K}. The points x and y either belong to the same K-complex or to two neighboring ones, and there is no loss of generality in assuming the former. By the argument above,

$$v(z)=-\alpha E^z\Big(\sum_{k=0}^{\sigma/\Delta t-1} v(Y_N(k\Delta t))\Delta t\Big)-\alpha E^z\Big(\sum_{k=\sigma/\Delta t}^{\infty} v(Y_N(k\Delta t))\Delta t\Big)$$

for $z=x,y$. As we now know that v is S-bounded, the first term is infinitesimal in both cases, and by (9.4) the ratio between the two second terms is almost one. It follows that $v(x)\approx v(y)$. The proof for $\tilde{A}_N$ is the same.

IX.13 Proposition: Assume that $N\in{}^*\mathbb{N}$ is infinite, that $t\in(0,\infty)$, and that $s\approx t$. Then

$$\text{trace}(T^t) = {}^\circ Tr_N(s)$$

The corresponding result holds for $\tilde{T}^t$.

Proof: Let $\alpha\in(-\infty,0]$ and let $\alpha_1,\alpha_2,..,\alpha_p$ be the eigenvalues of A_N that are infinitely close to α. Assume that each α_i has multiplicity m_i, and let $m=\sum_{i=1}^{p} m_i$. Since Lemma IX.9 tells us that the infinite eigenvalues do not contribute to $Tr_N(t)$, it suffices to show that $e^{\alpha t}$ is an eigenvalue of T^t with multiplicity m (- note that the proof also works when $p=0$!)

Let $v_1,v_2,..,v_m$ be an orthonormal set of eigenfunctions of A_N with eigenvalues $\alpha_1,\alpha_2,..\alpha_p$. By the proposition above, all the v_i are bounded and continuous, and hence they induce standard functions $\hat{v}_i:E\to\mathbb{R}$ in an obvious way. It's trivial to check that $\hat{v}_1,\hat{v}_2,..\hat{v}_m$ is an orthonormal set of eigenfunctions for T^t all with eigenvalue $e^{\alpha t}$.

It still remains to show that the multiplicity of $e^{\alpha t}$ isn't larger than m. If it were, I could find a function $\hat{v}\in L^2(E,m)$ orthonormal to $\hat{v}_1,\hat{v}_2,..,\hat{v}_m$ such that $T^t\hat{v}=e^{\alpha t}\hat{v}$. By elementary nonstandard measure theory, $\hat{v}$ has a lifting $v:F^{(N)}\to{}^*\mathbb{R}$ which is orthonormal to all the v_i and satisfies

$$\|Q_N^s v - e^{\alpha t} v\| \approx 0 \tag{9.5}$$

Let u be an eigenfunction of A_N corresponding to a finite eigenvalue β. If $\beta\approx\alpha$, then $(u,v)\approx 0$ by construction, and if $\beta\approx\alpha$, then on the one hand

$$(Q_N^s v,u) \approx e^{\alpha t}\,(v,u),$$

and on the other hand

$$(Q_N^s v,u) \approx (v,Q_N^s u) \approx e^{\beta t}(v,u)$$

which implies that $(v,u)\approx 0$. But this means that if $v=\Sigma c_i u_i$ is an eigenfunction expansion of v, then the contribution from the u_i corresponding to finite eigenvalues of A_N must be infinitesimal. This clearly contradicts (9.5) and the proof is complete.

All the necessary preparations have now been made, and we are ready to turn the heuristic argument at the end of Chapter III into a rigorous proof.

IX.14 Lemma. Assume that $t\in(0,\infty)$ and that $n\in \mathbb{N}_0$. Then

$$\text{trace}(T^t) \leq \mu^n \text{trace}(T^{\lambda^n t})$$

and

$$\text{trace}(\tilde{T}^t) \geq \mu^n \text{trace}(\tilde{T}^{\lambda^n t})$$

Proof: Choose an infinite $N\in {}^*\mathbb{N}$, and note that $F^{(N+n)}$ consists of μ^n N-complexes $F_1^{(N)}$, $F_2^{(N)}$,..,$F_{\mu^n}^{(N)}$ (I'm abusing my own terminology; strictly speaking each $F_i^{(N)}$ is an N-complex intersected with $F^{(N+n)}$). We have

$$\text{Tr}_{N+n}(t) = \sum_{x\in F^{(N+n)}} q_{N+n}^t(x,x) \leq \sum_{k=1}^{\mu^n} \sum_{x\in F_k^{(N)}} q_{N+n}^t(x,x),$$

with inequality because some elements belong to more than one complex. By construction of Y_{N+n} and Y_N,

$$\sum_{x\in F_k^{(N)}} q_{N+n}^t(x,x) \leq \sum_{y\in F^{(N)}} q_N^{t\cdot\lambda^n}(y,y),$$

with inequality because Y_{N+n} is usually not reflected when it hits a vertex in $F_k^{(N)}$. Hence

$$Tr_{N+n}(t) \leq \mu^n \sum_{y\in F^{(N)}} q^{t\cdot\lambda^n}(y,y) = \mu^n Tr_N(\lambda^n t),$$

and the first part of the lemma follows from the proposition above.

The proof of the second part is almost identical, but with all the inequalities reversed. Observe first that

$$\tilde{Tr}_{N+n}(t) = \sum_{x\in F^{(N+n)}} \tilde{q}^t_{N+n}(x,x) = \sum_{k=1}^{\mu^n} \sum_{x\in F_k^{(N)}} \tilde{q}^t_{N+n}(x,x)\sigma(x)^{-1},$$

where $\sigma(x)$ is the number of N-complexes to which x belongs. By construction of $\tilde{Y}_{N+n}$ and $\tilde{Y}_N$

$$\sum_{x\in F_k^{(N)}} \tilde{q}^t_{N+n}(x,x)\sigma(x)^{-1} \geq \sum_{y\in F^{(N)}} \tilde{q}_N^{t\lambda^n}(y,y),$$

where the inequality is due to the fact that $\tilde{Y}_{N+n}$ is not always killed when it hits a vertex in $F_k^{(N)}$. But then

$$\tilde{Tr}_{N+n}(t) \geq \mu^n \sum_{y\in F^{(N)}} \tilde{q}^{t\lambda^n}(y,y) = \mu^n \tilde{Tr}_N(\lambda^n t)$$

and the lemma follows from Proposition IX.13.

It is now easy to prove the following fundamental estimate.

IX.15 Proposition. There are constants $C,K\in \mathbb{R}_+$ such that

$$K\cdot t^{-\log\mu/\log\lambda} \leq \operatorname{trace}(\tilde{T}^t) \leq \operatorname{trace}(T^t) \leq C\cdot t^{-\log\mu/\log\lambda}$$

for all $t\in(0,1]$.

Proof: I only prove the inequality on the right, the one on the left follows from the same argument with all inequalities reversed.

Choose $C\in \mathbb{R}_+$ so large that

$$\text{trace}(T^t) \leq Ct^{-\log\mu/\log\lambda} \quad \text{for all} \quad t\in(\lambda^{-1},1],$$

and let

$$\zeta(t)=Ct^{-\log\mu/\log\lambda}.$$

Note that $\zeta(t)=\mu^n\zeta(\lambda^n t)$ for all positive t. Given $t\in(0,1]$, choose $n\in\mathbb{N}_0$ such that $\lambda^n t\in(\lambda^{-1},1]$. By the lemma,

$$\text{trace}(T^t) \leq \mu^n \text{trace}(T^{\lambda^n t}) \leq$$

$$\leq \mu^n\zeta(\lambda^n) = \zeta(t)=Ct^{-\log\mu/\log\lambda},$$

which is exactly what we wanted to prove.

We are now ready for the main theorem. Let

$n(\alpha)$=the number of eigenvalues of $-\Delta$ less than or equal to α counted with multiplicity.

IX.16 Theorem. There are constants $c,k\in\mathbb{R}_+$ such that

$$k\cdot\alpha^{\log\mu/\log\lambda} \leq n(\alpha) \leq c\cdot\alpha^{\log\mu/\log\lambda}$$

Proof: This follows immediately from Proposition IX.15 and a Tauberian theorem of the Karamata [19] type (see [10, Theorem 1]).

IX.17 Corollary. If d is the Hausdorff dimension of E, then

$$k\cdot\alpha^{d\log\nu/\log\lambda} \leq n(\alpha) \leq c\cdot\alpha^{d\log\nu/\log\lambda}. \tag{9.6}$$

As mentioned in the Introduction, Barlow and Perkins [5] have already proved this result for the Sierpinski gasket. Their proof is based on precise estimates on the transition probabilities, and gives better information about the size of the constants c and k. In situations were good estimates are hard to get, the sort of scaling argument used above may, however, be easier to generalize.

I pointed out in Chapter III that (9.6) can be regarded as a natural extension of Weyl's asymptotic formula

$$n(\alpha) \sim \alpha^{d \cdot \frac{1}{2}} \qquad \text{as} \quad \alpha \to \infty \tag{9.7}$$

for the Laplacian on a domain in $\mathbb{R}^d$ with finite volume. The trick is to consider the exponential factor $\frac{1}{2}$ in Weyl's formula as the exponent $\frac{1}{2}$ in the scaling law

$$(t,\omega) \to \alpha^{\frac{1}{2}} b\left(\frac{t}{\alpha}, \omega\right) \tag{9.8}$$

for ordinary Brownian motion. Since the scaling law for Brownian motion on a nested fractal is

$$(t,\omega) \to \alpha^{\log\nu/\log\lambda} B\left(\frac{t}{\alpha}, \omega\right), \tag{9.9}$$

we see how (9.6) generalizes (9.7). There is, however, an important difference between the two scaling laws; while (9.8) holds for all positive values of α, (9.9) only holds when α is an integer power of the time scaling factor λ. This difference is reflected in the distribution of eigenvalues; in the classical case the limit

$$\lim_{\alpha \to \infty} n(\alpha) \cdot \alpha^{-d/2}$$

exists (and is uniquely determined by the volume of the domain), but one can not expect the corresponding result to be true for fractals (see Barlow and Perkins [5] for more information about this and related topics).

X. Open problems.

In this paper, I have constructed Brownian motions on nested fractals and discussed a few of their simplest properties. Much remains to be done, and I end with a list of open problems of varying importance and difficulty:

1. Uniqueness: Is Brownian motion on a nested fractal unique? This is certainly the most important and also the most annoying open problem in the theory. If the answer is "yes", then proving that the map $\tilde{p}$ in Chapter V has a unique fixed point would be an important step toward a solution, but we should still have to take into account that the same fractal can be generated by different systems of similitudes. If the answer is "no", the situation is puzzling and a closer analysis will be necessary; it could, e.g., be the case that all but one fixed point of $\tilde{p}$ are unstable, and then the stable fixed point will give us the "natural" Brownian motion.

2. Remove the Nesting Axiom: If we remove the Nesting Axiom from our list of assumptions, all the fundamental ideas of this paper stop to make sense. To extend the theory in this direction seems very difficult, but it is also the most important challenge in trying to create a general theory for diffusions on fractals. The Sierpinski carpet seems a good testing-ground for new ideas; see Barlow and Bass [4].

3. Non-homogeneous scaling: A simpler way to generalize the theory is to allow the similitudes $\psi_1,..,\psi_\mu$ to have different scaling factors $r_1,..,r_\mu$, as mentioned in Chapter IV. And what happens if we allow them to vary randomly from one level to the next (see [12], [15]?

4. Domain of attraction: What is the domain of attraction for the fixed point(s) of the map $\tilde{p}$? Simple examples indicate that it is often very large and contains all points except obvious degeneracies; for the fractal in Figure 15, the only point not in the domain of attraction is the unstable fixed point $p_1=0$, $p_2=1$ (note that this point does not belong to $\mathfrak{P}$).

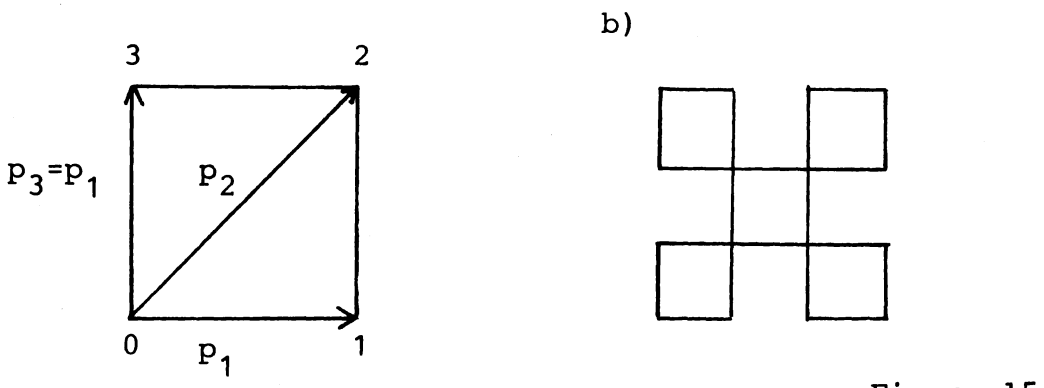


Figure 15

5. Scaling of time in the invariance principle: When is the number D_N in Theorem VIII.1 of order of magnitude λ^N?

6. Estimates on Green functions and transition probabilities: For the Sierpinski gasket, Barlow and Perkins [5] have wonderfully precise estimates; to what degree do they carry over to the general case?

7. Potential theory: As we saw in Chapter VII, the space of harmonic functions on E is finite dimensional. What happens on more complicated (sub-)domains? See Metz [28] for a study of potential theory on the Sierpinski gasket.

8. Transition times: Barlow and Perkins [5] use the theory of branching processes to study stable transition times on the Sierpinski gasket. Does their approach extend to nested fractals in general?

9. Modulus of continuity: What is the best value for the exponent β in Proposition VII.1 (see Barlow and Perkins [5] for the Sierpinski gasket)?

10. Point recurrence: Barlow and Perkins [5] have shown that Brownian motion on the Sierpinski gasket hits all points infinitely many times. Does this generalize to all nested fractals? (See also the next two questions).

11. Transience and recurrence: In this paper, I've been working with a compact state-space, but there is no problem in running the self-similarity process backwards to cover all of $\mathbb{R}^k$ with a fractal structure. When is Brownian motion on this extended fractal transient?

12. Local time: Barlow and Perkins [5] studied local time on the Sierpinski gasket. What happens in the general case?

13. Domain of the Laplacian: The paths of our Brownian motion B is normally of infinite quadratic variation, yet Fukushima's Decomposition Theorem [14, Theorem 5.2.2] tells us that if $u \in \mathcal{D}[\Delta]$, then $u(B)$ has finite quadratic variation! This shows that the elements in $\mathcal{D}[\Delta]$ must have rather strange properties; they must somehow be very smooth and very irregular at the same time! Normally, the only $\mathcal{C}^1$-functions in $\mathcal{D}[\Delta]$ are the constants (see Barlow and Perkins [5]). It would be very interesting to have a better description of the elements in $\mathcal{D}[\Delta]$; the next two problems outline possible approaches.

14. Intrinsic metric on E: Is there a metric on E which measures the distance "along" the fractal better than the Euclidean metric? Perhaps $\mathcal{D}[\Delta]$ could have a core of Lipschitz continuous elements with respect to such a metric?

15. Explicit construction of eigenfunctions: Is it possible to construct the eigenfunctions of Δ more or less explicitly by starting with the eigenfunctions of the discrete Laplacian on $F^{(1)}$, and then rescale them and patch them together in a systematic way?

16. Stochastic calculus: Is there a stochastic calculus associated with B? The answer has to be at least partially "yes" since there already is a stochastic calculus for Dirichlet forms (see Fukushima [14]), but one possible problem is that this theory only works for processes of the form $u(B_t)$ where $u \in \mathcal{D}[\sqrt{-\Delta}]$, and we have seen that only rather peculiar functions belong to this domain.

17. Eigenvalues of Schrödinger operators: Can the theory in Chapter IX be extended to operators of the form $-\Delta+V$? In particular, do "classical limit" considerations (see Simon [32, Section 10]) make sense on fractals?

18. Wave propagation: Study wave propagation on nested fractals. Kusuoka [21] has a section on the wave equation on the Sierpinski gasket, and Rammal and Toulouse [30] contains some interesting heuristic considerations.

19. Trace formula: The spectrum of the Laplace-Beltrami operator on the fundamental domain of a discrete group is described very precisely by Selberg's trace formula. Since self-similar fractals are reminiscent of fundamental domains in many respects, one may wonder whether there is a trace formula hiding somewhere.

20. Geometric construction of Brownian motion: Brownian motion on a domain E - be it Euclidean space, a manifold, or a fractal - is always intimately connected with the natural geometric measure on E. Is it possible to exploit this relationship to find a general method for constructing Brownian motions? If, for instance, we construct the Hausdorff measure on E by approximating E with cubic boxes of smaller and smaller size, is it possible to obtain Brownian motion on E as the limit of random walks between these boxes? To be really useful, such a general procedure would probably also have to guarantee the uniqueness of the limit process, and this seems to be a much harder problem.

References

1. S. Albeverio, J.E. Fenstad, R. Høegh-Krohn and T. Lindstrøm: Nonstandard Methods in Stochastic Analysis and Mathematical Physics, Academic Press, New York, 1986.

2. R.M. Anderson: A nonstandard representation for Brownian motion and Itô integration, Israel J. Math 25 (1976), 15-46.

3. J.R. Banavar and J.F. Willemsen: Probability density for diffusion on fractals, Phys.Rev. B 30 (1984), 6778-6779

4. M.T. Barlow and R.F. Bass: Construction of Brownian motion on the Sierpinski carpet, Preprint, 1987.

5. M.T. Barlow and E.A. Perkins: Brownian motion on the Sierpinski gasket, Preprint, 1987.

6. M.V. Berry: Distribution of modes in fractal resonators, in W. Güttinger and H. Eikemeier (eds.): Structural Stability in Physics, Springer-Verlag, Berlin, Heidelberg, New York, 1979, 51-53.

7. M.V. Berry: Some geometrical aspects of wave motion: wavefront dislocations, diffraction catastrophes, diffractals, in Geometry of the Laplace operator, Proc.Symp. Pure Math., Vol.36, Amer. Math.Soc., Providence, R.I., 1980, 13-38.

8. J. Brossard and R. Carmona: Can one hear the dimension of a fractal?, Commun. Math. Phys. 104 (1986), 103-122.

9. R.F. Brown: The Lefschetz Fixed Point Theorem, Scott, Foresman and Company, Glenview, Illinois, 1971.

10. L. De Hahn and U. Stadtmüller: Dominated variation and related concepts and Tauberian theorems for Laplace transforms, J. Math. Anal. Appl. 108 (1985), 344-365.

11. K.J. Falconer: The Geometry of Fractal Sets, Cambridge University Press, Cambridge, 1985.

12. K.J. Falconer: Random fractals, Camb. Phil. Soc. Math. Proc. 100 (1986), 559-582.

13. J. Feder: Fractals, Plenum Publishing Company, New York and London, 1988.

14. M. Fukushima: Dirichlet Forms and Markov Processes, North-Holland, Amsterdam, 1980.

15. S. Graf, R.D. Mauldin and S.C. Williams: The exact Hausdorff dimension in random recursive constructions, Mem. Amer. Math. Soc. 381, 1988.

16. S. Goldstein: Random walks and diffusions on fractals. Preprint, 1986.

7. R.A. Guyer: Diffusive motion on a fractal; $G_{nm}(t)$, Phys. Rev. A 32 (1985), 2324-2335.

18. J.E. Hutchinson: Fractals and self-similarity, Indiana Univ. Math. J. 30 (1981), 713-747.

19. J. Karamata: Neuer Beweis und Verallgemeinerung der Tauberschen Sätze, welche die Laplacesche und Stieltjes Transformation betreffen, J. Reine Angew. Math. 164 (1931), 27-39.

20. H. Kesten: Subdiffusive behavior of random walk on a random cluster, Ann. Inst. H. Poincaré 22 (1986), 425-487.

21. S. Kusuoka: A diffusion process on a fractal, in K. Itô and N. Ikeda (eds.): Probabilistic Methods in Mathematical Physics, Kino Kuniya - North Holland, 1987, 251-274.

22. M.L. Lapidus and J. Fleckinger-Pellé: Tambour fractal: vers une résolution de la conjecture de Weyl-Berry pour les valeurs propres du laplacien, C.R. Acad. Sci. Paris. Serie I, 306 (1988), 171-175.

23. T. Lindstrøm: Nonstandard analysis and perturbations of the Laplacian along Brownian paths, in S. Albeverio et al. (eds.) Stochastic Processes in Mathematics and Physics, LNM 1158, Springer-Verlag, 1986, 180-200.

24. T. Lindstrøm: Nonstandard energy forms and diffusions on manifolds and fractals, in S. Albeverio et al. (eds.): Stochastic Processes in Classical and Quantum Systems, Springer-Verlag, 1986, 363-380.

25. B.B. Mandelbrot: Les Objects Fractals: Forme et Dimension, Flammarion, Paris, 1975.

26. B.B. Mandelbrot: Fractals: Form, Chance, and Dimension, W.H. Freeman and Co., San Fransisco, 1977.

27. B.B. Mandelbrot: The Fractal Geometry of Nature, W.H. Freeman and Co., San Fransisco, 1983.

28. V. Metz: Brownsche Bewegung auf dem Sierpinski gasket aus potentialtheoretischer Sicht, Diplomarbeit, Bielefeld, 1988.

29. B.O'Shaughnessy and I. Procaccia: Diffusion on fractals, Phys. Rev. A 32 (1985), 3073-3083.

30. R. Rammal and G. Toulouse: Random walks on fractal structures and percolation clusters, J. Physique Lettres 44 (1983), L13-L22.

31. E. Seneta: Non-Negative Matrices, George Allen & Unwin Ltd, London, 1973.

32. B. Simon: Functional Integration and Quantum Physics, Academic Press, New York, 1979.

33. K.D. Stroyan and J.M Bayod: Foundations of Infinitesimal Stochastic Analysis, North-Holland, Amsterdam, 1986.

34. D. Williams: Diffusions, Markov Processes, and Martingales, John Wiley & Sons, Chichester, 1979.

SUBJECT INDEX

A

B

C

D

E

F

G

H

K

L

M

N

O

P

Q

R

S

T

U

V

W

Y

Tom Lindstrøm
Department of Mathematics
University of Oslo
P.O. Box 1053, Blindern
N-0316 Oslo 3
Norway